PARAMETRIC STATISTICAL INFERENCE

FOR

GEOMETRIC PROCESSES

by

SO-KUEN CHAN

A

Thesis

Submitted to

(Division of Statistics)

the Graduate School

of

The Chinese University of Hong Kong

In Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

(M. Phil.)

May, 1992

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled "Parametric Statistical Inference for Geometric Processes" submitted to the Graduate School by So-kuen Chan (陳素娟) in partial fulfillment of the requirement for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

Y. Lam

Dr. Y. Lam,
Supervisor

Lee Sik Yum

Dr. S.Y. Lee.

N.N. Chan

Dr. N.N. Chan

Professor L.J. Wei,
External Examiner

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

ACKNOWLEDGEMENT

The author of this thesis would like to express her deepest gratitude to her supervisor, Dr Yeh Lam who besides giving her invaluable advices and guidance, is always enthusiastic in giving every help and encouragement.

Besides, I would also like to give special thanks to the entire staff of the Department of Statistics. They are always eager to provide assistance whenever I need help.

Lastly, I have to thank my family for their everlasting support.

ABSTRACT

A stochastic process $\left\{ X_n, n=1,2,\ldots \right\}$ is a geometric process if there exists a > 0 called the ratio such that $\left\{ a^{n-1}X_n, n=1,2,\ldots \right\}$ is a sequence of IID random variables. This is a stochastically monotone process. Lam (1992b) studied the statistical inference for geometric processes by nonparametric method and used geometric process in modelling a point process with trend. In this thesis, under the assumption that X_1 follows one of the following lifetime distributions: Exponential, Gamma, Weibull and Lognormal distribution, we study the statistical inference for the geometric process by parametric method. The parameters a, λ and σ^2 where λ and σ^2 are respectively the mean and variance of X_1, are estimated by maximum likelihood method and modified maximum likelihood method.

CONTENT

CHAPTER ONE PREVIEW

Section 1 Introduction

When we buy a new car or a T.V. set, we would expect it to function properly for a reasonable period of time and then wear out gradually so that the consecutive operating times decrease and the consecutive repairing times increase until it breaks down completely.

If an item is likely to fail at any time after which a repair is provided to bring the item back to working condition, the operating time is a random variable with a distribution and we may model the series of operating times by a stationary point process or its corresponding counting process by a renewal process. If in addition, the successive interarrival times are independently, identically and exponentially distributed, we may model the data by a homogeneous Poisson process (HPP). [See Cox and Lewis (1966) and Ascher and Feingold (1984)]

However, due to the aging effect and cumulative wear, the item will gradually become more likely to fail and the failure rate increases with time. Data of this type, indeed most real life data, always follow a trend. One possible approach of describing this monotone trend is to use the nonhomogeneous Poisson process (NHPP) in which the hazard rate is monotone.

Lam (1988a,b and 1992a) first introduced the geometric process and studied its application to replacement problem. Then he (1992b) considered the statistical inference for geometric processes by nonparametric methods. In his paper (1992b), some simulation studies were performed so that suggestion regarding the best method under different conditions was made. Moreover, through the analysis of three sets of real data, a point process with trend can be satisfactorily modeled by a suitable geometric process.

Before going any further, we first give a simple definition of geometric process.

Definition

Given a sequence of random variables X_1, X_2, ... if for some $a > 0$, $\left\{ a^{n-1}X_n,\ n=1,2,\ldots \right\}$ forms a sequence of IID random variables, then $\left\{ X_n,\ n=1,2,\ldots \right\}$ is called a geometric process (GP) and a is called the ratio of the geometric process.

In fact, a geometric process is stochastically monotone and non-increasing if $a \ge 1$ and non-decreasing if $a \le 1$. If $a = 1$, the geometric process reduces to a sequence of stationary point process. variables.

For the geometric process $\left\{ X_n,\ n=1,2,\ldots \right\}$, we have the following results:

$$E(X_n) = \lambda/a^{n-1} \tag{1.1.1}$$

$$Var(X_n) = \sigma^2/a^{2(n-1)} \tag{1.1.2}$$

where λ and σ^2 are the mean and variance of X_1 respectively. Thus, a, λ and σ^2 completely determine the mean and variance of X_n. These three parameters are very important. For example, it is necessary to estimate a and λ in the application of geometric processes to the optimal replacement problem as shown in Lam (1988a,b).

In this thesis, by making an additional assumption that X_1 follows some distribution widely employed in describing the interarrival times or life-spans data such as the Exponential, Gamma, Weibull and Lognormal distributions, we use the parametric methods including the maximum likelihood estimation method and the modified maximum likelihood estimation method to estimate these three unknown parameters. Now, let us introduce the layout of this thesis.

In Chapter One, the preview chapter, we include in Section 1 a brief introduction which explain and outline the development of the geometric process. Then in Section 2, some distributions which are important in our context are introduced. Before any estimation procedures begin, some tests for geometric process are given in Section 3 and an outline of nonparametric inference from Lam (1992b) for the comparison with parametric inference is also given in this section. Finally, the test for distribution in parametric inference are listed in Section 4.

In Chapter Two, theories of estimation using maximum likelihood, modified maximum likelihood and modified moment, having Exponential, Gamma, Weibull and Lognormal distribution as the distributions of X_1 are devised.

Then simulation studies are performed and included in Chapter Three to evaluate the performance of each estimation method and to make suggestion of the choice of estimators under different situations.

In Chapter Four, some data sets are analyzed using the methodology developed in this thesis. Afterwards, in Chapter Five, the performance of each model in simulation and real data fitting using various nonparametric and parametric methods are analyzed and then in conclusion, suggestion and comment regarding the selection of the most appropriate estimator are made.

Section 2 The Life time Distribution

The statistical models for the distribution of X_1 or the sequence of IID random variables $\left\{ Y_n, n=1,2,... \right\}$ with $a^{n-1}X_n$ in the analysis of lifetime or interarrival time data are usually those which are positively skewed, originate from a finite threshold on the left and then tail off to zero on the right since it is rare to have a unusual long lifetime. We set the left threshold value to be zero in this thesis because the interarrival times can never be negative.

Apart from the distribution function F(t), the hazard rate function is also important for each distribution. It is defined as

$$h(t) = \frac{f(t)}{1 - F(t)} \qquad (1.2.1)$$

which is the instantaneous failure rate at time t given that the item has survived until time t. For example the successive operating time of a deteriorating system after repair will have an increasing hazard rate. Therefore, the hazard rate function is a factor of considerable importance in the selection of an appropriate model for describing a data set. When the shape parameter $\alpha > 1$, both the Gamma and Weibull distributions have an increasing hazard rate. This makes the Gamma and Weibull distributions applicable to many life testing experiments in which the 'aging effect' is expected. The Lognormal distribution would be an appropriate model when the failure rate is rather high initially and then decreases as the time t increases. A special case of Gamma or Weibull distribution is the Exponential distribution which has a constant failure rate and is thus suitable for another type of data which 'do not age'. Now, we first consider the Exponential distribution.

2.1 Exponential distribution

In life testing analysis, the simplest and most widely exploited model is the one-parameter Exponential distribution Exp(β) having density function

$$f(y|\beta) = \begin{cases} \beta e^{-\beta y} & y \geq 0 \\ 0 & y < 0. \end{cases} \tag{1.2.2}$$

Then

$$\lambda = E(Y) = \frac{1}{\beta}$$

and (1.2.3)

$$\sigma^2 = Var(Y) = \frac{1}{\beta^2}$$

where β is also called the scale parameter. Its density curve is reverse J-shaped. The constant hazard function is given by

$$h(y|\beta) = \beta. \tag{1.2.4}$$

It is therefore a suitable model for lifetime data where a used item is considered as good as new. This is the 'memoryless' property of the Exponential distribution.

Exponential distribution was widely used in early work, for example, in the reliability of the electronic components or in medical studies. Indeed, in many cases, the distribution of lifetime is either Exponential or can be approximated satisfactorily by an Exponential distribution. That is why Exponential distribution is important not only in theory but also in application, especially in lifetime testing and reliability theory.

Epstein (1958) remarked that the Exponential distribution played as important a role in life testing experiments as the part played by the Normal distribution in agricultural experiments on effects of different treatments in yield. [See Sinha (1986)]

2.2 Gamma distribution

We know that if α is an integer, Gamma distribution $\Gamma(\alpha,\beta)$ arises as the distribution of a sum of α independent and identically distributed random variables each having Exponential distribution $Exp(\beta)$ and its corresponding counting process will be Poisson distributed. The density function of the two parameters Gamma distribution $\Gamma(\alpha,\beta)$ is given by

$$f(y|\alpha,\beta) = \begin{cases} \frac{1}{\Gamma(\alpha)} \beta^{\alpha} y^{\alpha-1} \exp(-\beta y) & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases} \qquad (1.2.5)$$

where the shape parameter $\alpha > 0$, the scale parameter $\beta > 0$ and $\Gamma(\alpha)$ is the well-known Gamma function defined by

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt. \qquad (1.2.6)$$

Then

$$\lambda = E(Y) = \frac{\alpha}{\beta}$$

and (1.2.7)

$$\sigma^2 = Var(Y) = \frac{\alpha}{\beta^2}.$$

Depending on the skewness or the third standard moment τ_3 defined as

$$\tau_3(y) = E\{[y-E(y)]/\sigma\}^3 = 2/\sqrt{\alpha}, \tag{1.2.8}$$

the density curve can be either reverse J-shaped or bell-shaped. If $\tau_3 < 2$ or $\alpha > 1$, it is bell-shaped. If $\tau_3 \geq 2$ or $\alpha \leq 1$, it is reverse J-shaped. If $\tau_3 = 2$ or $\alpha = 1$, the Gamma distribution $\Gamma(1,\beta)$ reduces to Exponential distribution $Exp(\beta)$. [See Cohen and Whitten (1988)]

Note that there are no simple closed form expression available for the distribution function $F(y)$ and the hazard rate function $h(y)$. However, extensive studies shows that for $\alpha > 1$, $h(0) = 0$ and $h(y)$ approaches β asymptotically from below as $x \to \infty$. For $\alpha < 1$, $h(0) = \infty$ and $h(y)$ approaches β asymptotically from above as $x \to \infty$. This suggests that Gamma may be a useful model when items in a population are systems in a regular maintenance program. The failure rate may increase initially indicating the 'aging effect' but after some time the system would reach a stable condition due to maintenance and from then on, it would be as likely to fail in one time interval as in another. As $h(y) = \beta$, a constant when $\alpha = 1$, Gamma distribution also provides a generalization of Exponential distribution. [See Bain and Engelhardt (1991)]

2.3 Weibull distribution

Weibull distribution provides an alternate generalization of Exponential distribution. Suppose that instead of assuming that the failure time is distributed exponentially, we assume some power, say αth of the failure time is distributed exponentially. That is if $Z = Y^{\alpha}$ is distributed as $Exp(\beta)$, the random variable Y representing the failure time has Weibull distribution $W(\alpha,\beta)$ having density function

$$f(y|\alpha,\beta) = \begin{cases} \alpha\beta^{\alpha}\, y^{\alpha-1} \exp\{-(\beta y)^{\alpha}\} & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases} \tag{1.2.9}$$

where the shape parameter $\alpha > 0$ and the scale parameter $\beta > 0$.

Then

$$\lambda = E(Y) = \Gamma_1 / \beta$$

and (1.2.10)

$$\sigma^2 = \mathrm{Var}(Y) = [\ \Gamma_2 - \Gamma_1^{\ 2}\ ] / \beta^2$$

where $\Gamma_k = \Gamma(\ 1 + k/\alpha\ )$ and $\Gamma(y)$ is the Gamma function. It is sometimes expedient to make the following substitution

$$\theta = \beta^{\alpha} \quad \text{and} \quad \beta = \theta^{1/\alpha}. \tag{1.2.11}$$

Moreover, the hazard function is given by

$$h(y) = \alpha\beta^{\alpha}\, y^{\alpha-1} \tag{1.2.12}$$

and the third standard moment is given by

$$\tau_3(y) = \frac{\Gamma_3 - 3\Gamma_2\Gamma_1 + 2\Gamma_1^{\ 3}}{(\ \Gamma_2 - \Gamma_1^{\ 2}\ )^{3/2}}. \tag{1.2.13}$$

The density curve for this distribution can be either reverse J-shaped or bell-shaped depending on the shape parameter α or the skewness as measured by the third standard moment τ_3. As α becomes large, the density gets more peaked and symmetric around its mean, $\Gamma(1 + 1/\alpha) \simeq 1$. On the other hand, as θ decreases, the density gets less peaked and more asymmetric and has a rather long tail to the right. [See Sinha (1986)]

If $\tau_3 < 2$ or $\alpha > 1$, it is bell-shaped. As the hazard function h(y) is increasing, it becomes an appropriate model for items subjected to wear out. If $\tau_3 \geq 2$ or $\alpha \leq 1$, it is reversed J-shaped. As h(y) becomes decreasing, it might be an appropriate model for the item or system during its developmental stage when the elimination of problem sources results in increased reliability with the passage of time.

Exponential distribution with $\tau_3 = 2$ or $\alpha = 1$ and hence constant $h(y) = \beta$ is the special case of Weibull distribution. When $\alpha > \alpha_\circ = \mathbf{3.6023494257197}$ or $\tau_3 < 0$, it becomes negatively skewed. [See Cohen (1973)] When $\alpha = \alpha_\circ$ or $\tau_2 = 0$, it is almost normal in shape. Although τ_3 can be either positive, zero or negative, our primary interest lies in the case of positive skewness. In application concerning life-spans and interarrival times, values of α in excess of **3.22** or $\tau_3 < \mathbf{0.10}$ seldom occur. [See Cohen and Whitten (1988) and Bain and Engelhardt (1991)]

Weibull distribution has been named after the Swedish scientist Weibull who proposed it for the first time during an analysis of material strengths in 1939. [See Weibull (1939,1951)] It has been shown experimentally that this distribution provides a good fit for many different type of characteristics and is thus used extensively in life testing and reliability problems especially the 'wear-out' or fatigue failure such as the vacuum tube failures in Kao (1959) and the ball bearing failures in Lieblein and Zelen (1956). [See Sinha (1986)]

In some applications, there are also theoretical reasons for choosing Weibull model based on Extreme-value theory. As an example, suppose Y represents the strength of a chain of n links and let Y_i denotes the strength of the ith link. Then the strength of the chain is equal to the strength of its weakest link, i.e. $Y = \min(Y_i)$. Consequently the distribution of Y is the distribution of a minimum. For many different types of Y_i variables, the limiting distribution of the minimum approaches a Weibull distribution as $n \to \infty$. Thus, there is a direct relationship between the Weibull distribution and the type I Extreme-value distribution. If $Y \sim W(\alpha,\beta)$, then $Z = -\ln Y \sim EV(\phi,\theta)$ with $\theta = 1/\alpha$ and $\phi = \ln\beta$ and the distribution function is

$$F(z) = \exp\left[-\exp\left(- \frac{z-\phi}{\theta} \right) \right]. \tag{1.2.14}$$

Weibull distribution can also be related to a nonhomogeneous Poisson process with intensity $\nu(y) = (\alpha\beta)(y\beta)^{\alpha-1}$. This process will be referred as a Weibull process. [See Bain and Engelhardt (1991)]

2.4 Lognormal distribution

Lognormal distribution is related to Normal distribution in the same way as Weibull distribution related to Extreme-value distribution. If a random variable $X \sim N(\mu,\tau^2)$, then $Y = \exp(X) \sim LN(\mu,\tau^2)$. In the context of life testing and reliability problems, Lognormal distribution is preferred to Normal distribution since it ranges from 0 to $+\infty$ instead of $-\infty$ to $+\infty$. The density function of Lognormal distribution $LN(\mu,\tau^2)$ is given by

$$f(y|\mu,\tau^2) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,\tau y} \exp\left\{ -\dfrac{1}{2\tau^2} (\log y - \mu)^2 \right\} & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases} \qquad (1.2.15)$$

Sometimes, it is more convenience to make the following substitution

$$\beta = \exp(\mu) \quad \text{and} \quad \omega = \exp(\tau^2). \qquad (1.2.16)$$

Then

$$\lambda = E(Y) = \beta \sqrt{\omega}$$

and (1.2.17)

$$\sigma^2 = Var(Y) = \beta^2\omega(\omega-1).$$

Again, the hazard function h(y) cannot be expressed in closed form. However, it was proved that h(y) will be decreasing for large y. Thus, Lognormal distribution would be a good model when the failure rate is rather high initially and then decreases as time increases. For Normal distribution, the opposite is true: h(y) is increasing. Besides, while Normal distribution is symmetric about the mean μ, Lognormal distribution is quite skewed and all these aspects should be taken into consideration while choosing a model for the failure time distribution. [See Sinha (1986)]

Lognormal distribution arises in various contexts such as in physics (distribution of particles due to pulverization), economics (income distributions), biology (growth of organism) etc. Besides, it also has application in physical and industrial processes, textile research and quality control etc. [See Sinha (1986)]

2.5 Conclusion

Normal distribution is the limiting distribution of Gamma and Lognormal distributions as $\tau_3 \to 0$ but not of Weibull distribution. However, the standard Weibull distribution with $\tau_3 = 0$ is almost Normal with a left terminus at **-3.2431**. In addition, when $\tau_3 \le 1$, the discrepancies between Gamma, Weibull and Lognormal distributions are small except in the vicinity of the origins. For large τ_3, the differences are more pronounced. However, even for large τ_3, the right tails of these distributions tend to coincide. Distinguishing characteristics which might be important in model selection for a given data set are the mode, the threshold values, the skewness and the hazard function. [See Cohen and Whitten (1988)]

Section 3 Nonparametric Inference for Geometric Process (GP)

Suppose that we are given a data set of successive interarrival times $\left\{ X_n,\ n=1,2,\ldots,N \right\}$. Before any statistical analysis made, we have to test whether the given realization of point process confirms our definition of geometric process with ratio a. Then we may test whether there exists a trend in the point process. If there is no trend, we may further test whether the data confirm a stationary point process or its corresponding counting process confirm a renewal process. The following is the procedure of testing for a geometric process.

3.1 Test for Geometric Process (GP)

First of all, Laplace test is powerful in testing whether $\left\{ X_n,\ n=1,2,\ldots,N \right\}$ is a HPP process. Let us define the n^{th} arrival time T_n as

$$T_n = \sum_{i=1}^{n} X_i, \qquad n=1,2,\ldots,N.$$

Under HPP assumption,

$$U = \left\{ \frac{\sum_{n=1}^{N-1} T_n}{N-1} - \frac{T_N}{2} \right\} \Big/ \left\{ T_N \sqrt{\frac{1}{12(N-1)}} \right\} \sim N(0,1) \qquad (1.3.1)$$

asymptotically. [See e.g. Cox and Lewis (1966)]

To test the existence of a trend in successive X_n, many techniques, in particular, the graphical technique of plotting the cumulative number of events occurred against the cumulative time is an intuitive method of detecting the existence of a monotone trend. A concave graph shows that there is an increasing trend while a convex graph shows that there is a decreasing trend of the data.

If a trend probably exists, we test further whether the data come from a geometric process. Lam (1992b) suggested the following auxiliary series

$$U_m = X_{2m} / X_{2m-1}$$

and (1.3.2)

$$V_m = X_{m+1} X_{N+1-m} \quad \text{when N is odd}$$

$$V_m = X_m X_{N+1-m} \quad \text{when N is even,} \quad m = 1,2,\dots,M \qquad \text{where } M = \left[\frac{N}{2}\right]$$

so that all the information in the data set is utilized, no matter the size of the data set is odd or even. If $\left\{ X_n,\ n=1,2,\dots,N \right\}$ is a GP, U_m's and V_m's will be IID random variables. Now, let $\left\{ W_m,\ m=1,2,\dots M \right\}$ be such IID random variables and let I_A be the indicator of event A. Then the following two tests are useful in testing whether the data comes from a geometric process.

Turning points test (TP-test)

$$T_W = \sum_{m=2}^{M-1} I_{(\,(W_m - W_{m-1})(W_{m+1} - W_m) < 0\,)}, \qquad (1.3.3)$$

then

$$T_W^* = \left(T_W - \frac{2(M-2)}{3} \right) \Big/ \left(\frac{16M - 29}{90} \right)^{1/2} \sim N(0,1) \qquad (1.3.4)$$

asymptotically. [See Ascher and Feingold (1984)]

Difference - sign test (DS-test)

$$D_W = \sum_{m=2}^{M} I_{(W_m > W_{m-1})}, \tag{1.3.5}$$

then

$$D_W^* = \left(D_W - \frac{M-1}{2} \right) \Big/ \left(\frac{M+1}{12} \right)^{1/2} \sim N(0,1) \tag{1.3.6}$$

asymptotically. [See Ascher and Feingold (1984)]

Furthermore, whether the geometric process is stationary or stochastically monotone depends on the value of a. For this reason, Lam (1992b) suggested a nonparametric test based on linear regression technique. To start with, let us define

$$Y_n = a^{n-1} X_n, \qquad n = 1,2,\dots,N. \tag{1.3.7}$$

Then

$$\ln Y_n = (n-1) \ln a + \ln X_n, \qquad n = 1,2,\dots,N. \tag{1.3.8}$$

From the definition of geometric process, Y_n's are IID random variables and hence can be written as

$$\ln Y_n = \upsilon + e_n, \qquad n = 1,2,\dots,N. \tag{1.3.9}$$

where $E(\ln Y_n) = \upsilon$ and e_n's are also IID random variables with mean 0 and variance σ_e^2. Then (1.3.8) can be rewritten as

$$\ln X_n = (\upsilon + \ln a) - n \ln a + e_n, \qquad n = 1,2,\dots,N. \tag{1.3.10}$$

Therefore, a plot of $\ln X_n$ against n should exhibit a linear relationship. Besides, an upward trend may serve as an evidence of '$a < 1$' and a downward trend as an evidence of '$a > 1$' as shown in (1.3.10). In fact, (1.3.10) is a simple linear regression equation and the least square estimates (LSE) of the parameters are

$$\hat{\upsilon} = \frac{2}{N(N+1)} \left\{ (2N-1) \sum_{i=1}^{N} \ln X_i - 3 \sum_{i=1}^{N} (i-1) \ln X_i \right\} \tag{1.3.11}$$

$$\hat{\varphi} = \widehat{\ln a} = \frac{2}{(N-1)N(N+1)} \left\{ 3(N-1) \sum_{i=1}^{N} \ln X_i - 6 \sum_{i=1}^{N} (i-1) \ln X_i \right\} \tag{1.3.12}$$

and

$$\hat{\sigma}_e^2 = \left[\sum_{i=1}^{N} (\ln X_i)^2 - (\sum_{i=1}^{N} \ln X_i)^2 / N - \hat{\varphi} \left\{ (N-1) \sum_{i=1}^{N} \ln X_i / 2 - \sum_{i=1}^{N} (i-1) \ln X_i \right\} \right] \Big/ (N-2). \tag{1.3.13}$$

Furthermore, to test the testing hypothesis

$$H_o : a = 1 \qquad \text{against} \qquad H_1 : a \neq 1,$$

we can use the following test statistic

$$t = \frac{\hat{\varphi} \sqrt{(N-1)N(N+1)}}{\sqrt{12}\, \hat{\sigma}_e}. \tag{1.3.14}$$

Under H_o and the normality condition, $t \sim t(N-2)$. [See Lam (1992b) for details]

3.2 Nonparametric Estimation Method

In this section, we assume that a set of data $\left\{ X_n, n=1,2,\ldots,N \right\}$ is consistent with a geometric process and we estimate the parameters a, λ and σ^2 of the geometric process by the linear regression technique.

In fact, the plotting of $\ln X_n$ against n is not only useful for testing of geometric process but also provides a rough estimate of a [See examples in Chapter Four]. Moreover, the LSE of υ, lna and σ_e^2 are given by (1.3.11) to (1.3.13). Based on the LSE, Lam (1992b) gave the nonparametric estimates for a, λ and σ^2 as below.

From (1.3.12), it follows that the nonparametric estimation of a is given by

$$\hat{a}_{NP} = \exp(\hat{\varphi}). \tag{1.3.15}$$

In view of (1.3.7) and (1.3.9), we have

$$\begin{aligned}\sigma^2 &= \mathrm{Var}(X_1) = \mathrm{Var}(Y_n) = e^{2\upsilon}\,\mathrm{Var}[\exp(e_n)].\\ &\cong e^{2\upsilon}\,\mathrm{Var}(e_n) = e^{2\upsilon}\sigma_e^2\end{aligned} \tag{1.3.16}$$

and hence we can estimate σ^2 by

$$\hat{\sigma}_{NP1}^{\ 2} = \exp(2\hat{\upsilon})\ \hat{\sigma}_e^{\ 2}. \tag{1.3.17}$$

Moreover, as we define

$$\hat{Y}_n = \hat{a}^{n-1}\, X_n, \qquad n = 1,2,\ldots,N$$

where Y_n's are IID with variance σ^2, it is plausible to estimate σ^2 by

$$\hat{\sigma}_{NP2}{}^2 = \left\{ \sum_{i=1}^{N} \hat{Y}_i{}^2 - \left(\sum_{i=1}^{N} \hat{Y}_i \right)^2 \Big/ N \right\} \Big/ (N - 1) \tag{1.3.18}$$

the sample variance of $\hat{Y}_n$'s.

If a = 1, we can estimate σ^2 by sample variance

$$\hat{\sigma}_{NP3}{}^2 = \sum_{i=1}^{N} (X_i - \bar{X})^2 \Big/ (N - 1). \tag{1.3.19}$$

Furthermore, in view of the fact that $E(Y_n) = \lambda$, we can write

$$Y_n = \lambda(1 + \varepsilon_n) \tag{1.3.20}$$

with $E(\varepsilon_n) = 0$ and

$$Var(\varepsilon_n) = Var(Y_n) / \lambda^2 = \sigma^2 / \lambda^2. \tag{1.3.21}$$

It can be shown that λ must satisfy the following equation approximately

$$2\lambda^2 \ln\lambda - 2\upsilon\lambda^2 - \sigma^2 = 0. \tag{1.3.22}$$

Thus, we can estimate λ by solving the following equations

$$2\lambda^2 \ln\lambda - 2\upsilon\lambda^2 - \sigma_i{}^2 = 0 \qquad i = 1,2 \tag{1.3.23}$$

and get, say $\hat{\lambda}_{NP1}$ and $\hat{\lambda}_{NP2}$ respectively. In practice, the above two equations (1.3.23) can be solved numerically. Moreover, from (1.3.9),

$$\lambda = E(Y_n) = e^{\upsilon} E[\exp(e_n)]$$

$$= e^{\upsilon}(1 + \sigma_e^2 / 2). \tag{1.3.24}$$

Thus, the third estimate of λ is given by

$$\hat{\lambda}_{NP3} = (1 + \hat{\sigma}_e^2 / 2) \exp(\hat{\upsilon}) \tag{1.3.25}$$

To obtain another estimate of λ, we define $S_N = \sum_{i=1}^{N} X_i$. Then the other two estimates of λ are given by

$$\hat{\lambda}_{NP4} = S_N(1 - \hat{a}^{-1}) / (1 - \hat{a}^{-N}) \tag{1.3.26}$$

and

$$\hat{\lambda}_{NP5} = \sum_{i=1}^{N} \hat{Y}_i \Big/ N \tag{1.3.27}$$

respectively.

If a = 1, we can estimate λ by the sample mean

$$\hat{\lambda}_{NP6} = \bar{X} = \sum_{i=1}^{N} X_i \Big/ N. \tag{1.3.28}$$

Since X_n's are IID, $\hat{\lambda}_{NP6}$ and $\hat{\sigma}_{NP3}^2$ are obviously consistent. However, in general it is difficult to establish the consistency of the estimators for λ and σ^2.

Section 4 Test for Distribution

After all the necessary tests for geometric process, we then test whether the sequence of IID random variables $\left\{ Y_n, \; n=1,2,\ldots,N \right\}$ with $Y_n = a^{n-1}X_n$, $n = 1,2,\ldots,N$ confirms the given distribution.

4.1 Graphical method

Suppose that $\left\{ Y_n, \; n=1,2,\ldots,N \right\}$ follows some distributions, in particular, the Exponential distribution with distribution function

$$F(t) = 1 - e^{-\beta t}, \quad t \geq 0.$$

Then $-\ln\bar{F}(t) = \beta t$, $t \geq 0$. Again, if we form the order statistics

$$0 \leq Y_{(1)} \leq Y_{(2)} \leq \ldots\ldots \leq Y_{(N)}$$

and

$$\text{let } Y_{(0)} = 0 \quad \text{and} \quad Y_{(N+1)} = \infty,$$

the corresponding empirical distribution function (EDF) will be

$$F_N(t) = \frac{i}{N} \quad \text{for} \quad Y_{(i)} \leq t < Y_{(i+1)}, \quad i = 0,1,\ldots,N. \tag{1.4.1}$$

From Glivenko's theorem, we have

$$\sup_t \mid F_N(t) - F(t) \mid \xrightarrow[N\to\infty]{} 0. \tag{1.4.2}$$

Therefore, when N is large,

$$F(Y_{(i)}) \cong F_N(Y_{(i)}) = \frac{i}{N} \cong \frac{i}{N+1}. \tag{1.4.3}$$

Hence, the points

$$\left(Y_{(i)} \ , -\ln\left(1 - \frac{i}{N+1} \right) \right), \qquad i = 1,2,\ldots,N \tag{1.4.4}$$

will form approximately a straight line. [See Cao and Cheung (1986)]

On the other hand, if $\left\{ Y_n, \ n=1,2,\ldots,N \right\}$ follows Weibull distribution with distribution function

$$F(t) = 1 - \exp\{-(\beta t)^{\alpha}\},$$

then we may plot

$$\left((\beta Y_{(i)})^{\alpha} \ , -\ln\left(1 - \frac{i}{N+1} \right) \right). \tag{1.4.5}$$

In fact, for the other distributions such as Gamma and Lognormal, the graphical method can be devised in nearly the same way. However, due to the lack of closed form for the two distribution functions, we employ the usual way of plotting the points

$$\left(F(Y_{(i)}) \ , \frac{i}{N+1} \right), \qquad i = 1,2,\ldots,N. \tag{1.4.6}$$

It will form approximately a straight line if the proposed distribution of Gamma or Lognormal can satisfactorily model the given data set.

4.2 Kolmogorov-Smirnov Test (KS-test)

Let $F_N(y)$ be the empirical distribution function (EDF) of the sequence of IID random variables $\left\{ Y_n, \ n=1,2,\ldots,N \right\}$. Then we define two EDF statistics as

$$D^+ = \sup_y \{ F_N(y) - F(y) \} \tag{1.4.7}$$

and

$$D^- = \sup_y \{ F(y) - F_N(y) \}. \tag{1.4.8}$$

Another most well-known supremum EDF statistic, introduced by Kolmogorov (1933) is defined as

$$D = \sup_y \left| F_N(y) - F(y) \right| = \max (D^+, D^-) \tag{1.4.9}$$

and a closely related statistic V, given by Kuiper (1960) is defined as

$$V = D^+ + D^-. \tag{1.4.10}$$

Let $Y_{(1)} < Y_{(2)} < \ldots < Y_{(N)}$ be the order statistics for $\left\{ Y_n, \ n=1,2,\ldots,N \right\}$. By using the Probability Integral Transformation (PIT), $Z_i = F(Y_{(i)})$, $i=1,2,\ldots,N$ is uniformly distributed over $[0,1]$ and hence empirically

$$\hat{Z}_i = F_N(Y_{(i)}) \cong \frac{i}{N}, \qquad i=1,2,\ldots,N.$$

Then D can be rewritten as

$$D = \max_i \left\{ \left| \frac{i}{N} - F(Y_{(i)}) \right| , \left| F(Y_{(i)}) - \frac{i-1}{N} \right| \right\} = \max (D^+, D^-). \quad (1.4.11)$$

Another class of EDF statistics, the quadratic statistics given by the Cramér-von Mises family is defined as

$$Q = N \int_{-\infty}^{\infty} \left[F_N(y) - F(y) \right]^2 \psi(y) \, dF(y) \quad (1.4.12)$$

where $\psi(y)$ is a suitable function which gives weights to the squared difference $\left[F_N(y) - F(y) \right]^2$.

When $\psi(y) = 1$, Q is the Cramèr-von Mises statistic, usually denoted by W^2.

When $\psi(y) = \left[\{F(y)\}\{1 - F(y)\} \right]^{-1}$, Q is the Anderson-Darling (1954) statistic A^2.

A modification of W^2 is the Watson (1961) statistics U^2 which is defined as

$$U^2 = N \int_{-\infty}^{\infty} \left\{ F_N(y) - F(y) - \int_{-\infty}^{\infty} \left[F_N(y) - F(y) \right] dF(y) \right\}^2 dF(y). \quad (1.4.13)$$

Again, we form the order statistics of $Z_i = F(Y_{(i)})$,

$Z_{(1)} < Z_{(2)} < \ldots < Z_{(N)}$ and $\bar{Z} = \Sigma_i Z_i / N$.

Then we have

$$W^2 = \sum_{i=1}^{N} \left\{ Z_{(i)} - \frac{2i-1}{2N} \right\}^2 + \frac{1}{12N} \quad (1.4.14)$$

$$U^2 = W^2 - N(\bar{Z}-0.5)^2 \tag{1.4.15}$$

$$A^2 = -N - \left(\frac{1}{N}\right) \sum_{i=1}^{N} (2i-1) \left\{ \ln Z_{(i)} + \ln\left[1 - Z_{(N+1-i)} \right] \right\}. \tag{1.4.16}$$

The tabulated values for the Upper tail (UT) and/or Lower tail (LT) percentage point of some of $W(\alpha)$, $U(\alpha)$, $A(\alpha)$, $D^{+}(\alpha)$, $D^{-}(\alpha)$, $D(\alpha)$ and $V(\alpha)$ or their modifications are given separately for each distribution. Here α denotes the significance level. If say, $W^{*}(\alpha)_{LT} < W^{*} < W^{*}(\alpha)_{UT}$, we may accept from the null hypothesis that $\left\{ Y_n,\ n=1,2,\dots,N \right\}$ follows the given distribution.

For Exponential distribution, with β unknown and estimated by MLE, some modification of the EDF statistics are made

$$\begin{aligned}
W^{*} &= W^2(1.0 + 0.16/N) \\
U^{*} &= U^2(1.0 + 0.16/N) \\
A^{*} &= A^2(1.0 + 0.6/N) \\
D^{*} &= (D - 0.2/N)(\sqrt{N} + 0.26 + 0.5/\sqrt{N}) \\
V^{*} &= (V - 0.2/N)(\sqrt{N} + 0.24 + 0.35/\sqrt{N}).
\end{aligned} \tag{1.4.17}$$

The tabulated values for the Upper tail (UT) and Lower tail (LT) percentage point of $W^{*}(\alpha)$, $U^{*}(\alpha)$ and $A^{*}(\alpha)$ and the Upper tail (UT) percentage point of $D^{*}(\alpha)$ and $V^{*}(\alpha)$ are given in Table 1.4.1. [See Stephens (1974b, 1976a), Durbin (1975), Lilliefors (1969) and Margolin and Maurer (1976)]

For Gamma distribution, the parameters α and β can be estimated by solving the MLE likelihood equations

$$\left(\sum_{i=1}^{N} \ln Y_i \right) / N - \ln \bar{Y} = \Psi(\hat{\alpha}) - \ln \hat{\alpha}$$

and (1.4.18)

$$\hat{\beta} = \hat{\alpha} / \bar{Y}$$

where $\bar{Y} = \sum_{i=1}^{N} Y_i$ and $\Psi(x) = \dfrac{\partial \ln\Gamma(x)}{\partial x}$ is the psi function. The EDF statistics are calculated as (1.4.14) to (1.4.16) and no modification is required. The tabulated values for the Upper tail (UT) percentage point of $W(\alpha)$, $U(\alpha)$ and $A(\alpha)$ are given in Table 1.4.2. [See Lockhart and Stephens (1985b)]

For Weibull distribution, we may first make the transformation

$$Z_i = - \ln Y_i, \qquad i = 1,2,\ldots,N$$

so that Z_i will follow Extreme-value distribution $EV(\phi,\theta)$ with distribution function as (1.2.14) and $\theta = 1/\alpha$ and $\phi = \ln\beta$. Then we may test whether $\left\{ Z_i,\ i=1,2,\ldots,N \right\}$ follows Extreme-value distribution $EV(\phi,\theta)$ with the MLE estimates of parameters given by

$$\hat{\theta} = \sum_{i=1}^{N} Z_i \Big/ N - \left[\sum_{i=1}^{N} Z_i \exp(-Z_i / \hat{\theta}) \right] \Big/ \left[\sum_{i=1}^{N} \exp(-Z_i / \hat{\theta}) \right] \tag{1.4.19}$$

and

$$\hat{\phi} = - \hat{\theta} \ln \left[\sum_{i=1}^{N} \exp(-Z_i / \hat{\theta}) \Big/ N \right]. \tag{1.4.20}$$

(1.4.19) can be solved iteratively for $\hat{\theta}$ and then $\hat{\phi}$ can be determined in (1.4.20). By the invariance property of MLE, we also have

$$\hat{\theta} = 1/\hat{\alpha} \quad \text{and} \quad \hat{\phi} = \ln \hat{\beta}. \tag{1.4.21}$$

Again, some modifications for the EDF statistics should be made

$$\begin{aligned} W^* &= W^2(1.0 + 0.2/\sqrt{N}) \\ U^* &= U^2(1.0 + 0.2/\sqrt{N}) \\ A^* &= A^2(1.0 + 0.2/\sqrt{N}). \end{aligned} \tag{1.4.21}$$

The tabulated values for the Upper tail (UT) percentage point of $W^*(\alpha)$, $U^*(\alpha)$, $A^*(\alpha)$, $\sqrt{N}D^+(\alpha)$, $\sqrt{N}D^-(\alpha)$, $\sqrt{N}D(\alpha)$ and $\sqrt{N}V(\alpha)$ are given in Table 1.4.3. [See Stephens (1977) and Chandra, Singpurwalla and Stephens (1981)]

For Lognormal distribution, we should first make a transformation

$$Z_i = \ln Y_i, \qquad i=1,2,\ldots,N.$$

Then the MLE estimates for the mean μ and variance σ^2 of Normal distribution are the usual sample mean $\bar{Z}$ and variance $S_z^{\ 2}$ for the set $\left\{ Z_i,\ i=1,2,\ldots,N \right\}$. Again, some modifications for the EDF statistics should be made.

$$\begin{aligned} W^* &= W^2(1.0 + 0.5/N) \\ U^* &= U^2(1.0 + 0.5/N) \\ A^* &= A^2(1.0 + 0.75/N + 2.25/N^2) \end{aligned} \tag{1.4.23}$$

$$D^* = D^2(\sqrt{N} - 0.01 + 0.85/\sqrt{N})$$

$$V^* = V^2(\sqrt{N} + 0.05 + 0.82/\sqrt{N})$$

The tabulated values for the Upper tail (UT) and Lower tail (LT) percentage point of $W^*(\alpha)$, $U^*(\alpha)$ and $A^*(\alpha)$ and the Upper tail (UT) percentage point of $D^*(\alpha)$ and $V^*(\alpha)$ are given in Table 1.4.4. [See D'Agostino and Stephens (1986), Pearson and Hartley (1972) and Stephens (1974b)]

4.3 Chi-square Goodness-of-fit Test (χ^2 GOF-test)

This is perhaps the most popular goodness-of-fit test and is generally applicable for testing whether the given data set $\left\{ Y_n, n=1,2,\dots,N \right\}$ follows a proposed distribution F. Firstly, the sample range is divided into equal probability intervals

$$I_1 = (-\infty, a_1], \qquad I_2 = (a_1, a_2], \qquad \dots \qquad I_K = (a_{K-1}, \infty)$$

such that

$$p_i = p_j \text{ for } i \neq j$$

where $p_i = P(I_i) = F(a_i) - F(a_{i-1})$, $a_0 = -\infty$ and $a_K = \infty$. Then the expected number of observations in I_i from the sample with distribution F, is $E_i = Np_i$. To measure the disagreement between the expected E_i and the observed number of observation O_i in each I_i, we use the statistics

$$\chi^2 = \sum_{i=1}^{K} \frac{(O_i - E_i)^2}{O_i} \qquad (1.4.24)$$

Note that the number of intervals K should be carefully chosen so that K should be as large as possible with not more than 20% of O_i less than 5 and no O_i equals to 0. Under this condition,

$$\chi^2 \sim \chi^2(K - p - 1) \tag{1.4.25}$$

where p is the number of parameters estimated in the proposed distribution F.

4.4 F-test (Exponential distribution)

In practice, if a given set of data $\left\{ X_n,\ n=1,2,\ldots,N \right\}$ is consistent with a geometric process with $X_1 \sim \text{Exp}(\beta)$, $\left\{ Y_n,\ n=1,2,\ldots,N \right\}$ with $Y_n = a^{n-1}X_n$, n=1,2,...,N will generate a Poisson process. Let the order statistics be

$$Y_{(1)} \le Y_{(2)} \le \ldots\ldots \le Y_{(N)} \quad \text{and} \quad Y_{(0)} = 0.$$

Then the following test is useful

$$F = \frac{\sum_{i=1}^{K} D_i / K}{\sum_{i=K+1}^{N} D_i / (N-K)} \sim F_{2K,2(N-K)} \tag{1.4.26}$$

where $D_i = (N-i+1)(Y_{(i)} - Y_{(i-1)})$, $i = 1,2,\ldots,N$ and $K = \left[\frac{N}{2}\right]$.

[See Cao and Cheng (1986)]

CHAPTER TWO PARAMETRIC INFERENCE FOR GEOMETRIC PROCESS

Now, suppose that the given data set $\left\{ X_n, n=1,2,\ldots,N \right\}$ of the successive interarrival times is consistent with a geometric process with ratio a. Then $\left\{ Y_n, n=1,2,\ldots,N \right\}$ where

$$Y_n = a^{n-1}X_n, \qquad n=1,2, \ldots ,N \tag{2.1}$$

will be a sequence of IID random variables. Furthermore, we suppose that $\left\{ Y_n, n=1,2,\ldots,N \right\}$ follows the one of the following distributions: Exponential, Gamma, Weibull or Lognormal. Then in this section, we shall estimate the parameters a, λ and σ^2 of the geometric process using maximum likelihood and modified maximum likelihood method.

Note that as a is unknown, we estimate Y_n by substituting $\hat{a}$ for a in (2.1), that is $\hat{Y}_n = \hat{a}^{n-1} X_n$, n=1,2,...,N and we denote the two estimates of Y_n by $\hat{Y}_{NPn}$ and $\hat{Y}_{Pn}$ respectively if $\hat{a}$ substituted are respectively the nonparametric and parametric estimates. These $\hat{Y}_n$ are used for the true but unknown Y_i in the "lack of fit tests" introduced in the above section. As $\hat{a}$ is estimated from $\left\{ X_n, n=1,2,\ldots,N \right\}$, each $\hat{Y}_n$ is dependent and hence the assumptions of independent and identical distibuted Y_i for the tests are not fulfilled. We can only say that the tests are used approximately.

Section 1 Exponential distribution

If $\left\{ Y_n, n=1,2,\ldots,N \right\}$ follows Exponential distribution, that is $X_1 \sim \mathrm{Exp}(\beta)$, the density of X_n is given by

$$f_n(x_n) = a^{n-1} \beta \exp(-a^{n-1}\beta x_n). \tag{2.1.1}$$

Then the likelihood function becomes

$$L = a^{N(N-1)/2} \beta^N \exp\left\{ -\beta \sum_{i=1}^{N} a^{i-1} x_i \right\} \tag{2.1.2}$$

and

$$\frac{\partial \ln L}{\partial a} = \frac{N(N-1)}{2} \frac{1}{a} - \beta \sum_{i=2}^{N} (i-1) a^{i-2} X_i = 0 \tag{2.1.3}$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{N}{\beta} - \sum_{i=1}^{N} a^{i-1} X_i = 0. \tag{2.1.4}$$

Solving the above two equations, the maximum likelihood estimates (MLE) of a and β denoted by $\hat{a}_{E1}$ and $\hat{\beta}_{E1}$ respectively are given by

$$\sum_{i=1}^{N} (N-2i+1) \hat{a}_{E1}^{\ i-1} X_i = 0 \tag{2.1.5}$$

and $$\hat{\beta}_{E1} = \frac{N}{\sum_{i=1}^{N} \hat{a}_{E1}^{\ i-1} X_i}. \tag{2.1.6}$$

MLE of a in (2.1.5), $\hat{a}_{E1}$ can be solved by iterations. Then MLE of β, $\hat{\beta}_{E1}$ can be determined in (2.1.6). If we substitute a in (2.1.6) by the nonparametric estimate, we will obtain a modified MLE (MMLE) of $\hat{\beta}_{E2}$ given by

$$\hat{\beta}_{E2} = \frac{N}{\sum_{i=1}^{N} \hat{a}_{NP}^{\ i-1} X_i}. \tag{2.1.7}$$

Section 2 Gamma distribution

Assume that $X_1 \sim \Gamma(\alpha,\beta)$, then for all n, $X_n \sim \Gamma(\alpha, a^{n-1}\beta)$ with density

$$f_n(x) = \frac{(a^{n-1}\beta)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta a^{n-1}x). \tag{2.2.1}$$

Then the likelihood function becomes

$$L = \frac{a^{N(N-1)\alpha/2}\beta^{N\alpha}}{\Gamma(\alpha)^N} \left(\prod_{i=1}^{N} x_i\right)^{\alpha-1} \exp\left\{-\beta\sum_{i=1}^{N} a^{i-1}x_i\right\} \tag{2.2.2}$$

and

$$\frac{\partial \ln L}{\partial a} = \frac{N(N-1)}{2}\frac{\alpha}{a} - \beta\sum_{i=1}^{N}(i-1)\,a^{i-2}x_i = 0 \tag{2.2.3}$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{N(N-1)}{2}\ln a + N\ln\beta + \sum_{i=1}^{N}\ln x_i - N\,\Psi(\alpha) = 0 \tag{2.2.4}$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{N\alpha}{\beta} - \sum_{i=1}^{N} a^{i-1}x_i = 0 \tag{2.2.5}$$

where $\Psi(x) = \dfrac{\partial \ln\Gamma(x)}{\partial x}$ is the psi function.

Solving the above three equations, $\hat{a}_{G1}$, $\hat{\alpha}_{G1}$ and $\hat{\beta}_{G1}$ the maximum likelihood estimates (MLE) of a, α and β respectively are given by

$$\sum_{i=1}^{N}(N-2i+1)\,\hat{a}_{G1}^{\,i-1}X_i = 0 \tag{2.2.6}$$

$$\frac{(N-1)}{2}\ln\hat{a}_{G1} + \ln\hat{\alpha}_{G1} - \ln\left[\left(\sum_{i=1}^{N}\hat{a}_{G1}^{\,i-1}X_i\right)\Big/N\right] + \sum_{i=1}^{N}\ln X_i \Big/ N - \Psi(\hat{\alpha}_{G1}) = 0 \tag{2.2.7}$$

and

$$\hat{\beta}_{G1} = \frac{\hat{\alpha}_{G1} N}{\sum_{i=1}^{N} \hat{a}_{G1}^{\;i-1} X_i}. \tag{2.2.8}$$

MLE of a and α in (2.2.6) and (2.2.7) respectively can be solved by iteration. An approximation to $\hat{\alpha}_{G1}$ for used in the iteration process might be

$$\hat{\alpha} = \left[1 - N \Big/ \left(\bar{\hat{Y}}_P \sum_{i=1}^{N} \hat{Y}_{Pi}^{\;-1} \right) \right]^{-1} \tag{2.2.9}$$

where $\hat{Y}_{Pi} = \hat{a}_{G1}^{\;i-1} X_i$, n=1,2,...,N and $\bar{\hat{Y}}_P = \sum_{i=1}^{N} \hat{Y}_{Pi}$. [See Cohen and Whitten (1988)] Finally, MLE of β, $\hat{\beta}_{G1}$ can be determined in (2.2.8). Note that difficulties might be encountered when α is close to 1, even though it actually exceeds 1. Furthermore, the usual asymptotic properties of MLE do not hold unless $\alpha > 2$. Therefore, Johnson and Kotz (1970) recommended that MLE should be employed only if $\alpha > 2.5$ or the third standard moment $\tau_3 < 1.265$.

If we substitute the estimate of a in (2.2.7) to (2.2.8) by the non-parametric estimate $\hat{a}_{NP}$, we will get a modified MLE (MMLE) of $\hat{\alpha}_{G2}$ and $\hat{\beta}_{G2}$.

Since MLE are of doubtful utility when $\alpha \le 2.5$, the modified moment estimation (MME) method is suggested in that case since it is applicable over the entire parameter space and is unbiased with respect to both mean and variance. In this estimation method, we use $\hat{a}_{NP}$ for $\hat{a}$ and consider $\left\{ \hat{Y}_{NPn},\ n=1,2,\ldots,N \right\}$ in the deduction with the mean and variance respectively $\bar{Y}$ and S^2. For simplicity in notation, we write Y_n for $\hat{Y}_{NPn}$, n=1,2,...,N. Then MME are obtained by equating the first two sample moments to the corresponding distribution moments from (1.2.7) and we have

$$\hat{\alpha}_{G3} = (\bar{Y} / S)^2 \qquad (2.2.10)$$

and

$$\hat{\beta}_{G3} = \bar{Y} / S^2. \qquad (2.2.11)$$

Note that in the case that MLE is desired, MME will provide an excellent initial value in the iterative calculations of MLE and MME are applicable over all possible values of α.

Section 2 Weibull distribution

Assume that $X_1 \sim W(\alpha,\beta)$, the density function for X_n is given by

$$f_n(x_n) = a^{n-1}\, \alpha\theta\, (a^{n-1}x_n)^{\alpha-1} \exp[-\theta(a^{n-1}x_n)^{\alpha}] \qquad (2.3.1)$$

where $\theta = \beta^{\alpha}$ or $\beta = \theta^{1/\alpha}$. Then, the likelihood function becomes

$$L = a^{N(N-1)\alpha/2} (\alpha\theta)^N \left(\prod_{i=1}^{N} x_i \right)^{\alpha-1} \exp\left\{ -\theta \sum_{i=1}^{N} (a^{i-1}x_i)^{\alpha} \right\} \qquad (2.3.2)$$

and

$$\frac{\partial \ln L}{\partial a} = \frac{N(N-1)\alpha}{2a} - \frac{\alpha\theta}{a} \sum_{i=1}^{N} (i-1) (a^{i-1}x_i)^{\alpha} = 0 \qquad (2.3.3)$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{N(N-1)}{2} \ln a + \frac{N}{\alpha} + \sum_{i=1}^{N} \ln x_i - \theta \sum_{i=1}^{N} (a^{i-1}x_i)^{\alpha} \ln(a^{i-1}x_i) = 0 \qquad (2.3.4)$$

$$\frac{\partial \ln L}{\partial \theta} = \frac{N}{\theta} - \sum_{i=1}^{N} (a^{i-1} x_i)^{\alpha} = 0. \qquad (2.3.5)$$

Solving the above three equations, $\hat{a}_{W1}$, $\hat{\alpha}_{W1}$ and $\hat{\theta}_{W1}$ the maximum likelihood estimates (MLE) of a, α and θ respectively are the solution of the following equations

$$\sum_{i=1}^{N} (N-2i+1) (\hat{a}_{W1}^{\,i-1} X_i)^{\hat{\alpha}_{W1}} = 0 \qquad (2.3.6)$$

$$\frac{(N-1)}{2} \ln\hat{a}_{W1} + \frac{1}{\hat{\alpha}_{W1}} + \frac{\sum_{i=1}^{N} \ln X_i}{N} - \frac{\sum_{i=1}^{N} \left(\hat{a}_{W1}^{\,i-1} X_i \right)^{\hat{\alpha}_{W1}} \ln\left(\hat{a}_{W1}^{\,i-1} X_i \right)}{\sum_{i=1}^{N} \left(\hat{a}_{W1}^{\,i-1} X_i \right)^{\hat{\alpha}_{W1}}} = 0 \qquad (2.3.7)$$

and

$$\hat{\theta}_{W1} = \frac{N}{\sum_{i=1}^{N} (\hat{a}_{W1}^{\,i-1} X_i)^{\hat{\alpha}_{W1}}}. \qquad (2.3.8)$$

Note that MLE of a and α in (2.3.6) and (2.3.7) should be solved simultaneously for $\hat{a}_{W1}$ and $\hat{\alpha}_{W1}$. An IMSL Fortran program DNEQNF may be used for the numerical solution. Then MLE of θ, $\hat{\theta}_{W1}$ can be determined in (2.3.8). MLE of β is given by

$$\hat{\beta}_{W1} = \hat{\theta}_{W1}^{\,1/\hat{\alpha}_{W1}}. \qquad (2.3.9)$$

Note that computational problems are likely to be encountered when α is close to 1, even though it actually exceeds 1 because the distribution changes from reverse J-shaped to bell-shaped. Furthermore, the usual asymptotic properties of MLE do not hold unless $\alpha > 2$. Therefore, it is practically not considered advisable to employ MLE for $\alpha \le 2.2$. [See Cohen and Whitten (1988)]

If we substitute the estimate of a in (2.3.7) and (2.3.8) by the nonparametric estimate $\hat{a}_{NP}$, we will obtain a modified MLE (MMLE) of $\hat{\alpha}_{W2}$ and $\hat{\beta}_{W2}$.

Since MLE are of doubtful utility when $\alpha \le 2.2$, the modified moment estimation (MME) method is suggested in that case since it is applicable over the entire parameter space and is unbiased with respect to both mean and variance.

Again, we use $\hat{a}_{NP}$ for $\hat{a}$ in this estimation and consider $\left\{ \hat{Y}_{NPn},\ n=1,2,\ldots,N \right\}$ in the deduction of MME with respectively the mean and variance $\bar{Y}$ and S^2 . We also write Y_n for $\hat{Y}_{NPn}$, $n=1,2,\ldots,N$ for simplicity in notation. On equating the first two sample moments to corresponding distribution moments from (1.2.10), the estimating equations are

$$(\hat{\Gamma}_2 - \hat{\Gamma}_1^{\,2}) / \hat{\Gamma}_1^{\,2} = S^2 / \bar{Y}^2 \tag{2.3.10}$$

and

$$\hat{\beta}_{W3} = \hat{\Gamma}_1 / \bar{Y}. \tag{2.3.11}$$

Since (2.3.10) involves $\hat{\alpha}_{W3}$ as the only parameter, it can be solved iteratively. Then $\hat{\beta}_{W3}$ can be determined in (2.3.11). Although the calculation of MME requires considerably less computational effort than MLE, it should be remembered that estimate variances of the MLE are smaller than the corresponding variances of the MME. [See Cohen and Whitten (1988)]

Section 3 Lognormal distribution

Assume that $X_1 \sim LN(\mu,\tau^2)$, the density function for X_n is given by

$$f_n(x_n) = \frac{1}{\tau \sqrt{2\pi}\, x_n} \exp\left\{ - \frac{[\ \ln(a^{n-1}x_n) - \mu\ ]^2}{2\tau^2} \right\}. \quad (2.4.1)$$

Then the likelihood function becomes

$$L = \frac{1}{\tau^N (2\pi)^{N/2} \prod_{i=1}^{N} x_i} \exp\left\{ - \frac{\sum_{i=1}^{N} [\ \ln(a^{i-1}x_i) - \mu\ ]^2}{2\tau^2} \right\} \quad (2.4.2)$$

and

$$\frac{\partial \ln L}{\partial a} = - \frac{1}{a\tau^2} \sum_{i=1}^{N} [\ \ln(a^{i-1}x_i) - \mu\ ]\ (\ i - 1\ ) = 0 \quad (2.4.3)$$

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\tau^2} \sum_{i=1}^{N} [\ \ln(a^{i-1}x_i) - \mu\ ] = 0 \quad (2.4.4)$$

$$\frac{\partial \ln L}{\partial \tau} = - \frac{N}{\tau} + \frac{1}{\tau^3} \sum_{i=1}^{N} [\ \ln(a^{i-1}x_i) - \mu\ ]^2 = 0. \quad (2.4.5)$$

Solving the above three equations, $\hat{a}_{L1}$, $\hat{\mu}_{L1}$ and $\hat{\tau}_{L1}$ the maximum likelihood estimates (MLE) of a, μ and τ respectively are given by

$$\sum_{i=1}^{N} (\ N - 2i + 1\ ) \ln(\hat{a}_{L1}^{\ i-1}\ X_i) = 0 \quad (2.4.6)$$

$$\hat{\mu}_{L1} = \sum_{i=1}^{N} [\ \ln(\hat{a}_{L1}^{\ i-1}\ X_i)\ ] \Big/ N \quad (2.4.7)$$

and

$$\hat{\tau}_{L1}^{\ 2} = \sum_{i=1}^{N} \left[\ln(\hat{a}_{L1}^{\ i-1} x_i) - \hat{\mu}_{L1} \right]^2 \Big/ N. \tag{2.4.8}$$

Note that MLE of a, $\hat{a}_{L1}$ in (2.4.6) can be solved iteratively. Then $\hat{\mu}_{L1}$ and $\hat{\tau}_{L1}^{\ 2}$, the MLE of μ and τ^2 respectively, can be determined in (2.4.7) and (2.4.8) when $\hat{a}_{L1}$ is known. From (1.2.16), $\hat{\beta}_{L1}$ and $\hat{\omega}_{L1}$ the MLE of β and ω respectively, can be expressed as

$$\hat{\beta}_{L1} = \exp(\hat{\mu}_{L1}) \quad \text{and} \quad \hat{w}_{L1} = \exp(\hat{\tau}_{L1}^{\ 2}). \tag{2.4.9}$$

A modified MLE method (MMLE) is to substitute the estimate of a in (2.4.7) and (2.4.8) by the nonparametric estimate, $\hat{a}_{NP}$ to obtain a modified estimate of $\hat{\mu}_{L2}$ and $\hat{\tau}_{L2}$.

For MME, we again use $\hat{a}_{NP}$ for $\hat{a}$ and consider $\left\{ \hat{Y}_{NPn},\ n=1,2,\ldots,N \right\}$ in the deduction with mean and variance respectively $\overline{Y}$ and S^2. We also write Y_n for $\hat{Y}_{NPn}$, n=1,2,...,N for convenience. On equating the first two sample moments to the corresponding distribution moments from (1.2.17), the estimating equations are

$$\hat{\omega}_{L3} = (\ S\ /\ \overline{Y}\ )^2 + 1 \tag{2.4.10}$$

and

$$\hat{\beta}_{L3} = \overline{Y}\ /\ \hat{\omega}_{L3}^{\ 1/2} \tag{2.4.11}$$

The MME are unbiased with respect to the population mean and variance and they are also easy to calculate. Thus, they also provide an excellent first approximations in the iterative calculation of MLE if it is preferred.

Note that we may also substitute the estimate of a by 1 in all the ML equations and obtain another form of the modified MLE denoted by MLE1. Note that $\hat{\lambda}$ of the four distributions are denoted by respectively $\hat{\lambda}_{E3}$, $\hat{\lambda}_{G4}$, $\hat{\lambda}_{W4}$ and $\hat{\lambda}_{L4}$. This is the ordinary MLE for data without a trend. Care must be taken in verifying the assumption that a equals to 1. Nonparametric t-test [See Section 3.3.1 of Chapter One] could be employed.

Section 5 Remark

Computation problem is encountered in solving $\hat{\alpha}_{W1}$ and $\hat{a}_{W1}$ from (2.3.6) and (2.3.7) simultaneously. The estimates from DNEQNF are unsatisfactory. Hence, MLE of Weibull distribution are omitted. On the other hand, some estimates mentioned above can be shown to be equal. For a, we have

$$\hat{a}_{E1} = \hat{a}_{NP} \quad \text{and} \quad \hat{a}_{L1} = \hat{a}_{L2} \tag{2.5.1}$$

because respectively (2.1.5) & (2.2.6) are identical and (2.4.6) can be reduced to (1.3.12). For λ, we have

$$\hat{\lambda}_{NP5} = \hat{\lambda}_{E2} = \hat{\lambda}_{G2} = \hat{\lambda}_{G3} = \hat{\lambda}_{W3} = \hat{\lambda}_{L3} \tag{2.5.2}$$

$$\hat{\lambda}_{NP6} = \hat{\lambda}_{E3} = \hat{\lambda}_{G4} \tag{2.5.3}$$

and

$$\hat{\lambda}_{L1} = \hat{\lambda}_{L2} \quad \text{and} \quad \hat{\lambda}_{E1} = \hat{\lambda}_{G1} \tag{2.5.4}$$

because (2.5.2) and (2.5.3) equate $\hat{\lambda}$ to respectively $\bar{\hat{Y}}_{NP}$ and $\overline{X}$ while for (2.5.4), the a estimates are equal. For σ^2, we have

$$\hat{\sigma}_{NP2}^{\ 2} = \hat{\sigma}_{G3}^{\ 2} = \hat{\sigma}_{W3}^{\ 2} = \hat{\sigma}_{L3}^{\ 2} \tag{2.5.5}$$

since (2.5.5) equate $\hat{\sigma}^2$ to the sample variance of $\hat{Y}_{NPn}$, $S_Y^{\ 2}$.

CHAPTER THREE SIMULATIONS

Section 1 Introduction

In order to study the performance of different estimators for each distribution introduced in Chapter Two, we perform simulation experiment 300 realizations (I = 300) and in each realization we generate 201 data (N = 201) for the stationary point process $\left\{ Y_n,\ n=1,2,\ldots,201 \right\}$ which follows each of the four given distributions with its corresponding preassigned parameter values.

To do this, the International Mathematical and Statistical Library (IMSL) subroutine respectively DRNEXP for Exponential, DRNGAM for Gamma and DRNWIB for Weibull distribution are called to generate a sequence of IID random variables $\left\{ Z_n,\ n=1,2,\ldots,201 \right\}$ following the given distribution with the scale parameter β being set to one. Then by two steps transformation

$$Y_n = Z_n/\beta \tag{3.1.1}$$

and

$$X_n = Y_n/\ a^{n-1}, \tag{3.1.2}$$

the original realization will be transformed to the realization denoted by $\left\{ X_n,\ n=1,2,\ldots,201 \right\}$ which confirms a geometric process with ratio a and has the scale parameter β for its corresponding $\left\{ Y_n,\ n=1,2,\ldots,201 \right\}$.

However, for Lognormal distribution, the IMSL subroutine DRNLNL is called and the sequence of IID random variables $\left\{ Y_n,\ n=1,2,\ldots,201 \right\}$ so generated will follow $LN(\mu,\tau^2)$. Then by transformation (3.1.2) only, the realization $\left\{ X_n, n=1,2,\ldots,201 \right\}$ will also confirm a geometric process with ratio a.

By varying the values of the parameter set a, λ and σ^2, we obtain simulation results of each distribution. The values of a are set to be 0.90, 0.95, 1.00, 1.05 and 1.10 for each distribution because it is reasonable practically that a should be close to one. The values of λ for Exponential distribution and the values of $\lambda/\alpha/\mu$ and $\sigma^2/\beta/\tau^2$ for the other three distributions are given below

Exponential distribution :

$\lambda =$ 1.0, 2.0, 5.0 and 7.0.

Gamma distribution :

1. $\lambda = 1.2$ and $\sigma^2 = 1.0$
2. $\lambda = 1.6$ and $\sigma^2 = 1.0$
3. $\lambda = 2.0$ and $\sigma^2 = 1.0$
4. $\lambda = 1.4$ and $\sigma^2 = 1.2$
5. $\lambda = 1.8$ and $\sigma^2 = 1.2$.

Weibull distribution :

1. $\alpha = 6.0$ and $\beta = 6.0$
2. $\alpha = 6.0$ and $\beta = 10.0$
3. $\alpha = 8.0$ and $\beta = 4.0$
4. $\alpha = 12.0$ and $\beta = 6.0$
5. $\alpha = 16.0$ and $\beta = 8.0$.

Lognormal distribution :

1. $\mu = 0.4$ and $\tau^2 = 0.4$

2. $\mu = 0.5$ and $\tau^2 = 0.1$

3. $\mu = 0.5$ and $\tau^2 = 0.6$

4. $\mu = 1.0$ and $\tau^2 = 0.2$

5. $\mu = 2.0$ and $\tau^2 = 0.6$.

Note that in order to compare the performance of different estimation method in terms of its corresponding estimates, we use the same seed **1234576** each time for the generation of 300 sets of each 201 data and the results of these 300 realizations are averaged to give a single estimate.

Then for each realization $\left\{ X_n, n=1,2,\ldots,201 \right\}$, we test whether a trend exists by the tests given in Section 3.1 of Chapter One and we denote

$$P_U = P(\ |Z| \geq U\ ) \tag{3.1.3}$$

$$P_T^W = P(\ |Z| \geq T_W^*\ ) \tag{3.1.4}$$

$$P_D^W = P(\ |Z| \geq D_W^*\ ) \tag{3.1.5}$$

$$P_t = P(\ |Z| \geq t\ ) \tag{3.1.6}$$

where U, T_W^*, D_W^* and t are evaluated according to formulae (1.3.1), (1.3.4), (1.3.6) and (1.3.14) respectively and W's can be either U's or V's. Z is a standard Normal distributed random variable.

Note that the values of U, T_W^*, D_W^* and t and their standard derivation are averaged over each realization. Results show that these values are equal for each preassigned values of a except that P_U or P_t no longer takes the value zero when a = 1 and the point process reduce to a HPP because the TP-test and the DS-test utilize only the order information in each realization.

Afterwards, we have to test whether the fitted values $\left\{ \hat{Y}_n, n=1,2,\ldots,201 \right\}$ with $\hat{Y}_n = \hat{a}^{n-1} X_n$ follows the given distribution. Here a will be either the nonparametric estimate for MMLE & MME or the parametric estimate for MLE. From KS-test (Section 4.2 of Chapter One), values of W^2, U^2 and A^2 and/or $\sqrt{N}D^+$, $\sqrt{N}D^-$, $\sqrt{N}D$ and $\sqrt{N}V$ and sometimes their modifications will be compared with the table values to assess the significance of the test for different parameter values of a, λ and σ^2.

Besides, the Chi-square Goodness of Fit test (Section 4.3 of Chapter One) also measures the degree of confirmation of the fitted values $\left\{ \hat{Y}_n, n=1,2,\ldots,201 \right\}$ where $\hat{Y}_n = \hat{a}^{n-1} X_n$ with the given distribution. However due to the complexity of performing the test for each realization, we will not use this test here and we will reserve it later on in fitting the real data set in Chapter Four. For Exponential distribution, F-test is also available.

In order to compare the performance of different estimators, we define below three criteria to access the deviation of the estimates from the true value. Then we rank the estimators according to each of these criteria.

Intuitively, one criterion should measure the deviation of the estimate $\hat{\lambda}$ from the true λ. Then DM which is the average of deviation DM_i of the estimate $\hat{\lambda}_i$ from the true λ over each realization i, i=1,2,...,I is proposed. That is

$$DM = \sum_{i=1}^{I} DM_i \Big/ I \qquad \text{where } DM_i = | \lambda_i - \lambda |. \tag{3.1.7}$$

Alternatively, we may compare the observations X_{ni}, n = 1,2,...,N, that is the simulated values, with its fitted values $\hat{X}_{ni}$ to give the mean square error in each realization i, MSE_i , i = 1,2,...,I. Then MSE_i is averaged over i to give the MSE. That is

$$MSE = \left(\sum_{i=1}^{I} MSE_i \right) \Big/ I \qquad \text{where} \quad MSE_i = \sum_{n=1}^{N} (\hat{X}_{ni} - X_{ni})^2 / N \tag{3.1.8}$$

$$\text{and} \qquad \hat{X}_{ni} = \hat{\lambda}_i \, / \, \hat{a}_i^{\,n-1}$$

The first criterion which gives the average deviation of $\hat{\lambda}_i$ from the true λ is useful when our objective is parameter estimation while the second criterion which measures the deviation of the observations X_{ni}, n = 1,...,N from its fitted values $\hat{X}_{ni}$ in each realization i, i = 1,...,I is useful when our objective is data fitting. However, if our objectives are both parameter estimation and data fitting, a measure which is the sum of these two kind of criteria will be suitable,

$$MSEDM = \left(\sum_{i=1}^{I} MSEDM_i \right) \Big/ I \qquad \text{where} \qquad MSEDM_i = DM_i + \sqrt{MSE_i}. \tag{3.1.9}$$

For the comparison of estimate σ^2, the criterion directly followed from λ is again DMVAR which is the average of the deviation $DMVAR_i$ of the estimate $\hat{\sigma}_i^{\,2}$ from the true value σ^2 in each realization i, i=1,2,...,I. That is

$$DMVAR = \sum_{i=1}^{I} DMVAR_i \Big/ I \qquad \text{where} \qquad DMVAR_i = | \, \hat{\sigma}_i^{\,2} - \sigma^2 \, |. \tag{3.1.10}$$

Then we rank the estimates $\hat{\lambda}$ and $\hat{\sigma}^2$ according to these criteria. However, if the ranks of a particular estimation method are not consistent with different parameter set for given parameter a, the ranks are average. Estimate with the lowest average rank is considered the best estimate in comparison with the others.

Note that all the estimates given in the later sections are written in **six** decimal places whereas the standard deviation of the corresponding estimates in bracket and the results of various test are written in **four** significant figures or decimal places. Estimates that are equal up to six decimal places are considered equal and are usually listed once only. For simplicity in notation, we write Tables 3.1-4.1 to stand for Tables 3.1.1, 3.2.1, 3.3.1 and 3.4.1 for respectively Exponential, Gamma, Weibull and Lognormal distribution.

Section 2 Simulation Results

For the result of simulations, the values of P_U, P_T^U, P_D^U, P_T^V and P_D^V in the test for a trend are given in Tables 3.1-4.1 of the four distributions : all except P_U have values close to 0.5 and they are independent of a but dependent on the parameter values except Exponential distribution. For P_U, it is always 0.0000 when $a \neq 1$ but also close to 0.5 when $a = 1$ for all the four distributions.

Then in the test for GP, the p-values of the t-test of whether $a = 1$ is always as low as 0.0000 when $a \neq 1$ but are 0.4800, 0.4783 and 0.4977 for respectively Exponential, Weibull and Lognormal distribution when $a = 1$. For Gamma distribution, the p-values are given in Table 3.2.1b. They are also 0.0000 when $a \neq 1$ and although they depend on the parameter values when $a = 1$, they are all very close to 0.5. The results of these two tests of process type perfectly support theoretical expectation for each distribution.

Then for the test of distribution, the values of W^2, U^2 and A^2 and/or $\sqrt{N}D^+$, $\sqrt{N}D^-$, $\sqrt{N}D$ and $\sqrt{N}V$ and sometimes their modifications are given in Tables 3.1-4.2 of the four distributions. Note that the results are independent on the parameter a except MLE1 which is applicable only when $a = 1$. However, the results are also dependent on the parameter values of λ and σ^2 for Gamma and Weibull distribution. But for Exponential distribution, the values of EDF statistics are independent of the parameter λ and for Lognormal distribution, the values of EDF statistics are even independent of the estimation method MLE, MMLE and MME. When compared with the critical values of Tables 1.4.1 - 1.4.4, the results also justify the theoretical conclusion. For Exponential distribution, another test used is the F-test. The p-value is 0.5337 for MLE and 0.5331 for MMLE which are again equal for different parameters a, λ and σ^2. For MLE1, the p-value is 0.5453 when $a = 1$. These results again confirm the theoretical expectation.

After all the necessary tests of process type and distribution, we can estimate the parameters accordingly by the methodology developed in Chapter Two.

The estimates of parameters a, λ, σ^2 and β / α&β / μ&τ^2 are given in Tables 3.1-4.3 to Tables 3.1-4.6 respectively of the four distributions. In Tables 3.1-4.3, the estimates of a using parametric method (MLE) and nonparametric method (NP) are listed. Results show that for Exponential distribution, the estimates of a for different set of parameters λ and σ^2 are equal while for Gamma distribution, the estimates of a from either MLE or NP depend on the parameters λ and σ^2. However, for reasons mentioned in the remark of Chapter Two, only NP estimation of a are listed for Weibull and Lognormal distribution.

Then the various parametric and nonparametric estimates of λ are given in Tables 3.1-4.4 of the four distributions. Note that these estimates are equal for different values of a except NP4 and NP6 & MLE1 of each distribution give meaningful estimates only when $a = 1$.

The estimates of σ^2 are given in Tables 3.1-4.5 of the four distributions. Under the assumption that X_n's are IID, NP3 and MLE1 give meaningful estimates of σ^2 only when $a = 1$. Note that the estimates are again equal for different values of a.

Lastly, the parameter α for Exponential distribution, parameters α and β for Gamma and Weibull distribution and parameters μ and τ^2 for Lognormal distribution arising from the parametric method of estimation are given in Tables 3.1-4.6 of the four distributions. Note that the estimates are again equal for different values of a except MLE1 which is again applicable only when $a = 1$.

Note that in fact, the MLE of each of λ, σ^2, α and β of Gamma distribution are subject to small variation for different values of a but this is not reflected in values written up to 6 or sometimes 5 decimal places. For example, when $a = 1.10$, MLE of σ^2 is 1.200249 instead of the tabled value 1.200248 for the preassigned parameters $\lambda = 1.4$ and $\sigma^2 = 1.2$. Besides, the MLE and MMLE of λ and μ of Lognormal distribution and the MMLE of each of λ, σ^2, α and β of Weibull distribution when $\alpha = 6$ are also subject to small variation in the 6 decimal place for $a > 1$. In the case that the estimates are slightly varied in the 6 decimal places for different a, we list only the value of estimate when $a = 1.00$ for the sake of simplicity.

In order to compare the performance of parametric and nonparametric estimation, we rank the various parametric and nonparametric estimates of $\hat{\lambda}$ and $\hat{\sigma}^2$ according to the criteria developed above. The results of $\hat{a}$, $\hat{\lambda}$ and $\hat{\sigma}^2$ are given separately in the next section. For Exponential distribution, the ranks of each estimation method given the parameter a are the same for different parameter λ except the ranks with stars which are the average of different ranks for different λ. However, for the remaining three distributions, the ranks are not always the same for given λ and estimation method. Then the average ranks for each a and estimation method are given. Results of these ranking are given in the seventh table of each distribution. Since the ranks sometimes vary for given a, λ and estimation method especially Lognormal distribution, the choice of a good estimator for λ becomes rather complicated. Thus, we will continue to discuss the performance of these estimates in more detail in the following section.

Section 3 Comment and Discussion

From the results of comparison, we give a brief comment over the performance of different parametric estimation method for each distribution.

1. Parameter a

For Exponential distribution, Table 3.1.3 shows that the parametric estimate is always better since it has value closer to the true value than the nonparametric estimate but for Gamma distribution, the relative performance of these two estimates depends on parameters λ and σ^2. From Table 3.1.3 and 3.2.3, the parametric and nonparametric estimates of a are almost the same. Therefore, it is reasonable to replace a by $\hat{a}_{NP}$ in MMLE and MME for the estimation of other parameters. For Weibull distribution, exact MLE is difficult and we will use mainly MMLE or MME. For Lognormal distribution, the parametric and nonparametric estimates are equal.

2. Parameter λ

The result is rather complicated: it depends on the value of a and the criteria using since the performance of the estimates in view of estimation (DM) or data fitting (MSE) or both (MSEDM) are rather different. Note that the first choice of estimator according to MSEDM always follows that of MSE when $a < 1$ and DM when $a > 1$ and $a = 1$. This is reasonable because when $a < 1$, the MSE becomes very large and hence we should place more importance in data fitting but when $a > 1$, the MSE becomes much smaller and estimation will be relatively more important. The results are summarized in Table 3.5.1.

3. Parameter σ^2

In contrary to λ, the ranking of the estimates $\hat{\sigma}^2$ according to DMVAR are always equal. When $a \neq 1$, the first choice of $\hat{\sigma}^2$ is MLE and when $a = 1$, the first choice of $\hat{\sigma}^2$ is MLE1.

CHAPTER FOUR EXAMPLES

The methodology developed in Chapter Two is used to analyze six data set. Three of the data set are the same as those studied by Lam (1992b) while the remaining three are taken from [Cox and Lewis (1966)]. For each data set, we perform the test for Geometric Process and a plot of $\ln X_n$ against n is also given to see whether a trend exists in the data set. Since we estimate the unknown parameters using parametric method besides the nonparametric method, we also perform test for distribution for the fitted values $\left\{ \hat{Y}_n,\ n=1,2,\ldots,N \right\}$ after fitting the suggested parametric models to check whether the models are appropriate in the sense that the fitted values follow that distribution. Then we make comparison between the performance of different method of estimation. Again for simplicity in notation, we write Tables 4.1-6.1 to stand for Tables 4.1.1, 4.2.1, 4.3.1, 4.4.1, 4.5.1 and 4.6.1 for respectively the data set Coal mining disasters, Air1, No3, Air2, No4 and Patients.

For the six data sets, the results of the test for GP is shown in Tables 4.1-6.1. Then the tests for distribution which include KS-test, F-test and χ^2 GOF-test are shown in Tables 4.1-6.2. Note that the mark '•' in each box of the table indicates that the null hypothesis for the distribution is not rejected at all the three given significance level α : 0.025, 0.05 or 0.10 while the mark '⊙' indicates that whether the null hypothesis is rejected depends on the value of α chosen. Note also that in the χ^2 GOF test, the degree of freedom for the test is shown in bracket. It varies sometimes even for a particular data set at a particular distribution because once the conditions of the test are not satisfied [See Section 4 of Chapter One], the number of intervals chosen will be automatically decreased by one and then the test repeats until the chosen minimum number of intervals, 5 is reached and the process stops.

Finally, for the comparison of the performance of different estimators, we rank them according to MSE and estimators having the same MSE are given the same rank. Results of different estimates and their ranking are shown in Tables 4.1-6.3 of the six data sets. Then for the purpose of data fitting, we choose estimator having the lowest MSE. We found that the results are almost in perfect agreement with the results from the simulation studies as shown in Table 3.5.1.

However, the performance of estimation of parameters for different estimators can not be measured by the criterion DM in (3.1.7) for real data set because the parameters if exist are not known. In this case, we may follow the suggestions from simulation studies. We will demonstrate this in the following examples. Note that the estimators chosen for the purpose of estimation may not have the lowest MSE but should have its MSE quite similar to the lowest MSE. Therefore, we propose one criterion for measuring the relative size of MSE in addition to the relative rank in evaluating the performance of different estimators. This is a ratio of MSE defined as

$$MSER = [\ MSE - \min(MSE_i)\ ] \ / \ \min(MSE_i) \qquad (5.1)$$

where i stands for different estimators. When we choose estimators, we should also take into consideration other estimators that though not rank the first according to MSE, have MSE only slightly greater as shown from MSER. In order to see how the chosen parametric and/or nonparametric models fit the data, a plot of real data and fitted values using the chosen estimators for data fitting and estimation are given for each data set.

Note that the notation, for example 'MLE_E' in Tables 4.1-6.3 of the six data sets stands for MLE of Exponential distribution.

4.1 Coal mining disasters data

The data consists of 190 intervals in days between successive disasters in Great Britain. It is originally given by Maguire et al. (1952) and were thoroughly studied by Cox and Lewis (1966) as an example in the analysis of trend. Later on, the data were corrected and extended to cover the period 15th May 1851 to 22nd March 1962 inclusive, a total of 40550 days. [See Andrews and Herzberg (1985)]. The data set contains a zero because there has been two accidents on the same day. This zero is being replaced by 0.5 since usually two accidents would not occur simultaneously but instead, were separated by some time interval say half a day. [See Jarrett (1979)]

From Table 4.1.1, the data are consistent with a geometric process with ratio a not equal to 1. Again, the plot of $\ln X_n$ against n in Graph 4.1.1 as suggested in Lam (1992b) shows that a linear relationship between them exists and thus the data do come from a GP. In addition, the plot seems to exhibit a small rising trend which implies that $a < 1$ from (1.3.10). This is again confirmed by the fact that the estimates $\hat{a}$ using either parametric or nonparametric method have value less than but very close to 1.

From Table 4.1.2, Exponential and Gamma distribution with estimators MLE and MMLE and Weibull distribution generally fit the data. For data fitting, NP4 is the best estimator of λ from Table 4.1.3 and from the simulation studies of Lam (1992b), the nonparametric estimate of $\hat{\sigma}^2$, NP2 should also be chosen.

For estimation, we follow the results of the simulation studies when $a < 1$ and consider only the suggested estimators of λ for the distributions which fit the data as shown in Table 3.5.1, i.e. MLE of Exponential and Gamma distribution and MMLE of Weibull distribution. Then MLE of Exponential distribution is chosen since although its rank according to MSE is 9, the difference of this MSE from the lowest MSE is small as shown from 0.17% of MSER. A plot of real data and the fitted values of NP4 and MLE of Exponential distribution is given in Graph 4.1.2.

4.2 Air1 data

This data set together with the Air2 data in the latter section is taken from Cox and Lewis (1966). They are the intervals in operating hours between successive failures of air conditioning equipment from third and sixth air crafts respectively [See Proschan (1963)]. The Air1 data consist of 29 data while the Air2 data consist of 30 data.

From Table 4.2.1, the data are consistent with a GP. The p-values, P_u and P_t for the Air1 data shows that it is in border line that $a \neq 1$. However, estimates of a using either parametric or nonparametric method have values less than 1. This shows that the data exhibit a small increasing trend as shown also in the plot of $\ln X_n$ against n in Graph 4.2.1.

Although it passes nearly all the test for distribution, the small size of the data set makes the power of the tests limited. For data fitting, NP3 is chosen for $\hat{\lambda}$ and thus for $\hat{\sigma}^2$, NP2 should be chosen.

For estimation when $a < 1$, we consider all the suggested estimators of λ for each distribution from the simulation studies, i.e. MLE of Exponential and Gamma distribution, MMLE of Weibull distribution and NP3. Then again NP3 is chosen since although it has the lowest rank according to MSE. A plot of real data and the fitted values of NP3/NP4 is given in Graph 4.2.2.

4.2 No3 data

This data set consists of the times of unscheduled maintenance actions for U.S.S. Halfbeak No.3 main propulsion diesel engine in hours. It was studied by Ascher and Feingold (1969 and 1981) and in Lam's paper (1992b), a geometric process was used to fit this data set satisfactorily. After discarding the arrival times to scheduled engine overhauls since we are only interested in the arrival times of failures which cause the unscheduled maintenance actions, the number of data becomes 71.

From Table 4.3.1, the data are strongly consistent with a GP with ratio a not equal to 1. Besides, the plot of $\ln X_n$ against n in Graph 4.3.1 shows that a linear relationship exhibits with a small rising trend which implies that $a > 1$. The various estimates of a also support this conclusion.

From Table 4.3.2, Gamma and Weibull distribution seem to describe the data well. Again, for data fitting, NP4 is chosen for $\hat{\lambda}$ and thus for $\hat{\sigma}^2$, NP2 should be chosen.

For estimation when $a > 1$, we consider the suggested estimators of λ from the simulation studies for the distributions which fit the data, i.e. the MLE of Gamma distributions and MMLE of Weibull distribution. Then MMLE of Weibull distribution is chosen since although having rank 2 according to MSE, its MSE is only slightly greater than the lowest MSE as shown from 0.01% of MSER. A plot of real data and the fitted values of NP4 and MMLE of

Weibull distribution is given in Graph 4.3.2.

4.4 Air2 data

This data set consisting of 30 data is also taken from Cox and Lewis (1966). From Table 4.4.1, the data are consistent with a GP. However, the p-values of P_u and P_t show that it is just significance that $a \neq 1$. In addition, the estimates $\hat{a}$ using either nonparametric or parametric methods have values larger than 1. This shows that the data exhibit a small decreasing trend as shown in the plot of $\ln X_n$ against n in Graph 4.4.1.

Although it also passes nearly all the tests for distribution for $a \neq 1$, the power of the test is limited for the small data size. For data fitting, NP4 is the best estimator for $\hat{\lambda}$ and for $\hat{\sigma}^2$, NP2 should be chosen.

For estimation when $a > 1$, we consider all the suggested estimators of λ from the simulation studies for each distribution, i.e. MLE of Exponential and Gamma distribution, MMLE of Weibull distribution and NP3. Then MMLE of Weibull distribution is chosen since although having rank 3 according to MSE, its MSE is only slightly greater than the lowest MSE as shown from 0.02% of MSER. A plot of real data and the fitted values of NP4 and MMLE of Weibull distribution is given in Graph 4.4.2.

4.5 No4 data

This data set contains the arrival times to unscheduled maintenance actions for the U.S.S. Grampus No.4 main propulsion diesel engine in hours. It was tabulated and studied by Lee (1980 a,b). Lam (1992b) has also analyzed these data. There is an extremely large interarrival time 6930 which by further study, was revealed that "the person who recorded failures went on leave and nobody took his place until his return" [See Ascher and Feingold (1984)]. Therefore, this outlier and its successor 575 should be cast away. Again, the times of scheduled engine overhauls are also discarded as in No3 data. Besides, the time interval between two unscheduled maintenance actions taken in the same hour is replaced by half hour as in Coal disaster data. There are 56 data afterwards.

From Table 4.5.1, this data are consistent with a HPP. The plot of $\ln X_n$ against n in Graph 4.5.1 shows no apparent trend. All these are evidence that a might be equal to 1.

From Table 4.5.2, generally all except Weibull when $a = 1$ and Lognormal distribution seem to describe the data well. For data fitting, MLE of Exponential or Gamma distribution will be the best choice of estimator for the parameters.

For estimation when $a = 1$, we consider the suggested estimators of λ from the simulation studies for the distributions which fit the data, i.e. MLE1 of Exponential and Gamma distribution. Then MLE1 of Exponential distribution or NP6 is chosen since although having rank 3 according to MSE, its MSE is only slightly greater than the lowest MSE as shown from 1.71% of MSER. A plot of real data and the fitted values of MLE and MLE1 of Exponential distribution is given in Graph 4.5.2.

4.6 Patients data

This data set is also taken from Cox and Lewis (1966). It is the intervals in hours between successive arrival of patients at a intensive care unit. The original data are the arrival times consisting of year, month, day, hour and minutes. Since both the 28th, 31st and 214th arrival times are identical to the proceeding arrival times, we have their corresponding interarrival times 0 replaced by 0.04 hr (≈ 2.5 mins). This is because the arrival times are recorded up to 5 mins and for reasons as in Coal disaster data, half this recording interval are used. Besides, it may due to recording error that the 247th arrival times is earlier than the proceeding one. For this reasons, the 247th arrival time and onwards are truncated and thus we have altogether 245 data of interarrival times.

From Table 4.6.1, this data are consistent with a HPP. The plot of $\ln X_n$ against n in Graph 4.6.1 shows no apparent trend. All these are evidence that a might be equal to 1.

From Table 4.6.2, generally no particular distribution seem to describe the data well. Maybe on the contrary, the large size of this data set makes the tests too sensitive to departure from hypothesis. Further research on other distributions should be conducted before suggestion on estimators made. However, if best possible estimators of parameters should be chosen now, MLE of Exponential or Gamma distribution might be the best choice for data fitting.

For estimation when $a = 1$, we consider all the suggested estimators of λ from the simulation studies, i.e. MLE1 for each distribution. Then MLE1 of Exponential or Gamma distribution is chose. Since it has rank 13 according to MSE with MSER 3.43%, it may not be a good estimator for the purpose of data fitting. A plot of real data and the fitted values of MLE and MLE1 of Exponential distribution is given in Graph 4.6.2.

CHAPTER FIVE COMPARISON AND CONCLUSION

For simulation, the ranking of estimates $\hat{\lambda}$ and $\hat{\sigma}^2$ for different a are quite consistent for Exponential distribution and thus we have some confidence in asserting their performance in spite of the fact that the parameter values of λ chosen for the simulation are not plentiful. For other distributions, due to the inconsistence and complexity of ranking especially Lognormal distribution, further simulation works are necessary to verify these results.

In employing various estimates, care should be taken. Although MLE1 has a better performance of estimation over other estimates when $a = 1$ as shown from the result of simulation studies, we should test this underlying assumption first before applying the method. Otherwise, other estimates using MLE or NP should be used. Moreover, regularity problem are likely encountered in employing MLE or MMLE of Gamma or Weibull distribution when they are in the vicinity of the transition point $\alpha = 1$ where the shape changes from bell-shaped to reverse J-shaped. However, although Johnson and Kotz [See Johnson and Kotz (1970)] suggested that α should be greater than 2.5 for Gamma distribution, MLE gives satisfactorily estimates for α as low as 1. For α less than 1, computational problems are encountered. Thus, MME is particularly useful in estimation when the shape parameter α is less than 2.5 and MLE or MMLE becomes inapplicable. [See Cohen and Whitten (1988)] However, for Lognormal distribution, there is no parameter limit for MLE.

For nonparametric estimation, while NP4 have a rather good performance in terms of data fitting, the performance of NP1 and NP3 are rather unsatisfactorily except Lognormal distribution. Besides, as the ranking of parametric estimates of Lognormal distribution are in general rather unsatisfactory, the use of parametric estimates of this distribution is discouraged. In conclusion, Lognormal distribution seems to be quite different from the Exponential, Gamma and Weibull family.

In the contrary to simulation, the performance of the parametric estimates for real data set is not satisfactory enough according to MSE. This may due to the simple fact that the underlying assumption of the distribution is not satisfied. In fact there are many life-time distributions besides the four distributions discussed in this thesis, for example, Inverse Gaussian and Extreme-value distributions which are also of considerable importance in life span model. Therefore, if the performance of parametric estimation is not satisfactory enough, it may due to the fact that right model of distribution should be the one apart from the four distributions Exponential, Gamma, Weibull and Lognormal distributions and further research is necessary to study these parametric inference.

Besides, from the results of simulation and real examples, we find that the nonparametric estimators in particular NP4, always have small MSE. This is because LSE is employed in the deduction of nonparametric estimates. Thus they are particularly useful in data fitting while their performance in estimating λ is usually not good enough when compared with the parametric estimates. For real data, since it is impossible to evaluate the performance of estimation by measuring the deviation from the true parameters because the true parameters if exist are forever unknown. Thus we have only one criteria, the MSE to consider in the evaluation of the performance of each estimators.

Furthermore, although results of KS-test, F-test and χ^2 GOF-test agree one another closely, the power of these tests for distribution is affected by the size of data set. We find that data set of small size generally gives insignificant results except perhaps very extreme case while large data set generally gives significant results. The discrimination power of these tests against different distributions is not great enough. This is another difficulties facing in the evaluation of the fitness of the data set to the distribution model.

In conclusion, when a data set is given, we first use the tests for GP to see whether it satisfies the assumption of a GP. If the result is negative, we may stop and conclude that our method of parametric estimation is inapplicable for this data set. If the result is positive, we may further test whether a is less than, equal to or greater than 1. For this purpose, the plotting of $\ln X_n$ against n is also useful.

Then if the distribution of the data set is known to be one of the four distributions we study, we can choose estimators according to the suggestions from simulation studies in Table 3.5.1. However, if the distribution of the data set is not known, we should first check the data for the mode, the threshold values, the skewness and the hazard function etc. to see if they possess the distinguishing characteristics of a particular distribution. This gives valuable insight in choosing a suitable model for the data set. Finally if all these information is still insufficient to make a judgment with confidence, we should choose estimators of λ following the suggestions from simulation studies according to whether $a = 1$ and whether data fitting or estimation or both is/are desired. Then estimators of λ with the smallest MSE should be chosen for data fitting. If our objective is estimation, we may choose the estimators of λ among all the suggested estimators with MSER as low as possible so that this MSE is not significantly greater than the lowest MSE. The suitability of the chosen model for the data set is then evaluated by the tests for distribution using the fitted values $\left\{ \hat{Y}_n,\ n=1,2,\ldots,N \right\}$ under that model.

TABLES AND GRAPHS

Table 1.4.1 : Test for Exponential Distribution Using EDF Statistics.

Sign.	W^*		U^*		A^*		D^*	V^*
α	LT	UT	LT	UT	LT	UT	UT	UT
0.025	0.0233	0.271	0.0207	0.189	0.178	1.591	1.184	1.774
0.050	0.0276	0.222	0.0243	0.159	0.208	1.321	1.094	1.655
0.100	0.0338	0.175	0.0293	0.129	0.249	1.062	0.995	1.527

Table 1.4.2 : Test for Gamma Distribution Using EDF Statistics.

Est. $\hat{\alpha}$	Significance level α								
	0.10	0.05	0.025	0.10	0.05	0.025	0.10	0.05	0.025
	W^2 (UT)			U^2 (UT)			A^2 (UT)		
1	0.111	0.136	0.162	0.098	0.119	0.141	0.657	0.786	0.917
2	0.107	0.131	0.155	0.097	0.118	0.139	0.643	0.768	0.894
3	0.106	0.129	0.153	0.097	0.118	0.138	0.639	0.762	0.886
4	0.105	0.128	0.152	0.097	0.117	0.138	0.637	0.759	0.883
5	0.105	0.128	0.151	0.097	0.117	0.138	0.635	0.758	0.881
6	0.105	0.128	0.151	0.097	0.117	0.138	0.635	0.757	0.880

Table 1.4.3 : Test for Weibull Distribution Using EDF Statistics.

Sign.	UT										
α	W^*	U^*	A^*	$\sqrt{N}D^+$		$\sqrt{N}D^-$		$\sqrt{N}D$		$\sqrt{N}V$	
				N>50	N=50	N>50	N=50	N>50	N=50	N>50	N=50
.100	.102	.097	.637	.734	.727	.733	.724	.803	.790	1.372	1.344
.050	.124	.117	.757	.808	.796	.808	.796	.874	.856	1.477	1.453
.025	.146	.138	.877	.877	.870	.877	.860	.939	.922	1.557	1.538

Table 1.4.4 : Test for Normal Distribution Using EDF Statistics.

Sign.	W^*		U^*		A^*		D^*	V^*
α	UT	LT	UT	LT	UT	LT	UT	UT
0.025	0.148	0.019	0.136	0.018	0.873	0.139	0.995	1.585
0.050	0.126	0.022	0.117	0.021	0.752	0.160	0.895	1.489
0.100	0.104	0.026	0.096	0.025	0.631	0.188	0.819	1.386

Table 3.1.1 : Test for Geometric Process for Exponential Distribution.

a	Is it a HPP ?		Is it a GP ?							
	P_U	S.D.	P_T^U	S.D.	P_D^U	S.D.	P_T^V	S.D.	P_D^V	S.D.
a≠1	.0000	.0000	.4888	.0164	.4861	.0162	.5071	.0164	.4978	.0164
a=1	.4932	.0159	.4888	.0164	.4861	.0162	.5071	.0164	.4978	.0164

Table 3.1.2 : Test for Exponential Distribution Using EDF Statistic.

Method	W^*	U^*	A^*	D^*	V^*
MLE	0.0843	0.0674	0.5547	0.0020	0.0038
MMLE	0.0846	0.0675	0.5575	0.0019	0.0037
MLE1	0.0830	0.0665	0.5472	0.0020	0.0036

Table 3.1.3 : Estimation of Parameter a for Exponential Distibution.

Method	a				
	0.90	0.95	1.00	1.05	1.10
MLE	0.900121 (6.092E-5)	0.950128 (6.431E-5)	1.000134 (6.769E-5)	1.050141 (7.108E-5)	1.100148 (7.446E-5)
NP	0.900127 (7.926E-5)	0.950134 (8.366E-5)	1.000141 (8.806E-5)	1.050148 (9.247E-5)	1.100155 (9.687E-5)

Table 3.1.4 : Estimation of Parameter λ for Exponential Distribution.

Method	λ = 1.00 $\left(\begin{array}{l}\sigma^2=1.00\\ \beta=1.00\end{array}\right)$	2.00 $\left(\begin{array}{l}\sigma^2=4.00\\ \beta=0.50\end{array}\right)$	5.00 $\left(\begin{array}{l}\sigma^2=25.0\\ \beta=0.20\end{array}\right)$	7.00 $\left(\begin{array}{l}\sigma^2=49.0\\ \beta=1/7\end{array}\right)$
MLE	1.014173 (0.0082)	2.028345 (0.0164)	5.070863 (0.0410)	7.099209 (0.0573)
MMLE / NP5	1.022378 (0.0105)	2.044755 (0.0209)	5.111888 (0.0523)	7.156644 (0.0732)
MLE1 / NP6 (a=1.00)	0.995926 (0.0041)	1.991852 (0.0083)	4.979631 (0.0207)	6.971483 (0.0289)
NP1	0.845314 (0.0089)	1.690622 (0.0177)	4.226550 (0.0444)	5.917169 (0.0621)
NP2	0.983905 (0.0103)	1.967802 (0.0206)	4.919494 (0.0514)	6.887288 (0.0720)
NP3	1.051706 (0.0110)	2.103412 (0.0220)	5.258530 (0.0551)	6.887288 (0.0772)
NP4 (a=0.90)	1.027990 (0.0180)	2.055980 (0.0360)	5.139950 (0.0901)	7.195930 (0.1261)
(a=0.95)	1.023936 (0.0150)	2.047873 (0.0300)	5.119681 (0.0751)	7.167554 (0.1051)
(a=1.00)	1.018910 (0.0103)	2.037821 (0.0207)	5.094552 (0.0517)	7.132372 (0.0723)
(a=1.05)	1.005983 (0.0096)	2.011967 (0.0193)	5.029916 (0.0482)	7.041883 (0.0674)
(a=1.10)	0.999685 (0.0128)	1.999370 (0.0255)	4.998425 (0.0638)	6.997795 (0.0893)

Table 3.1.5 : Estimation of Parameter σ^2 for Exponential Distribution.

Method	σ^2			
	1.00 $\begin{pmatrix}\lambda=1.00\\ \beta=1.00\end{pmatrix}$	4.00 $\begin{pmatrix}\lambda=2.00\\ \beta=0.50\end{pmatrix}$	25.00 $\begin{pmatrix}\lambda=5.00\\ \beta=0.20\end{pmatrix}$	49.00 $\begin{pmatrix}\lambda=7.00\\ \beta=1/7\end{pmatrix}$
MLE	1.048601 (0.0171)	4.194404 (0.0683)	26.215027 (0.4267)	51.381454 (0.8363)
MMLE	1.077988 (0.0224)	4.311952 (0.0895)	26.949698 (0.5592)	52.821408 (1.0959)
MLE1 (a=1.00)	0.996983 (0.0083)	3.987931 (0.0331)	24.924567 (0.2066)	48.852152 (0.4050)
NP1	0.562889 (0.0121)	2.251558 (0.0482)	14.072237 (0.3013)	27.581584 (0.5905)
NP2	1.069721 (0.0245)	4.278884 (0.0979)	26.743025 (0.6118)	52.416328 (1.1992)
NP3 (a=1.00)	0.983340 (0.0111)	3.933359 (0.0446)	24.583495 (0.2786)	48.183649 (0.5461)

Table 3.1.6 : Estimation of Parameter β for Exponential Distribution.

Method	β			
	1.00 $\begin{pmatrix}\lambda=1.00\\ \sigma^2=1.00\end{pmatrix}$	0.50 $\begin{pmatrix}\lambda=2.00\\ \sigma^2=4.00\end{pmatrix}$	0.20 $\begin{pmatrix}\lambda=5.00\\ \sigma^2=25.0\end{pmatrix}$	0.142857 $\begin{pmatrix}\lambda=7.00\\ \sigma^2=49.0\end{pmatrix}$
MLE	1.005346 (8.130E-3)	0.502673 (4.065E-3)	0.201069 (1.626E-3)	0.143621 (1.162E-3)
MMLE	1.008749 (1.025E-2)	0.504375 (5.123E-3)	0.201750 (2.049E-3)	0.144107 (1.464E-3)
MLE1 (a=1.00)	1.009349 (4.265E-3)	0.504674 (2.132E-3)	0.201870 (8.529E-4)	0.144193 (6.092E-4)

Table 3.1.7 : Comparison of Estimates using Parametric and Nonparametric method for Exponential Distribution.

a	$\hat{\lambda}$	DM	MSE	MSEDM	$\hat{\sigma}^2$	DMVAR
0.90	NP1	5	6	6	NP1	4
	NP2	2	5	3	NP2	3
	NP3	4	4	5	MMLE	2
	NP4	6	1	1	MLE	1
	MMLE / NP5	3	3	4		
	MLE	1	2	2		
0.95	NP1	5	6	6	NP1	4
	NP2	2	3.75 *	3	NP2	3
	NP3	4	5	5	MMLE	2
	NP4	6	1	1	MLE	1
	MMLE / NP5	3	3.25 *	4		
	MLE	1	2	2		
1.00	NP1	7	7	7	NP1	6
	NP2	3	5	5	NP2	5
	NP3	6	6	6	NP3	2
	NP4	4	2	4	MMLE	4
	MMLE / NP5	5	3	5	MLE	3
	MLE	2	1	2	MLE1	1
	MLE1 / NP6	1	4	1		
1.05	NP1	6	6	6	NP1	4
	NP2	3	3	3	NP2	3
	NP3	5	5	5	MMLE	2
	NP4	2	1	2	MLE	1
	MMLE / NP5	4	4	4		
	MLE	1	2	1		
1.10	NP1	6	6	6	NP1	4
	NP2	2	3	2	NP2	3
	NP3	4	5	4	MMLE	2
	NP4	5	1	5	MLE	1
	MMLE / NP5	3	4	3		
	MLE	1	2	1		

Table 3.2.1a : Test for Geometric Process for Gamma Distribution.

Est.		Is it a HPP ?		Is it a GP ?							
λ	σ^2	P_U	S.D.	P_T^U	S.D.	P_D^U	S.D.	P_T^V	S.D.	P_D^V	S.D
		a≠1									
		.0000	.0000	Same as below							
		a=1									
1.2	1.0	.5453	.0156	.5026	.0167	.5017	.0164	.4918	.0167	.5143	.0162
1.6	1.0	.6537	.0135	.5159	.0163	.4821	.0164	.4998	.0162	.5244	.0164
2.0	1.0	.7011	.0126	.5165	.0158	.5033	.0163	.4946	.0170	.5022	.0163
1.4	1.2	.5999	.0150	.5027	.0167	.4828	.0168	.4830	.0163	.5017	.0168
1.8	1.2	.6614	.0136	.5048	.0166	.4904	.0162	.4759	.0164	.5085	.0161

Table 3.2.2 : Test for Gamma Distribution Using EDF Statistic.

Estimates		Method	W^*	U^*	A^*
λ	σ^2				
1.2	1.0	MLE	0.0650	0.0592	0.4041
		MMLE	0.0649	0.0591	0.4031
		MME	0.0791	0.0756	0.5584
		MLE1 (a=1)	0.0662	0.0603	0.4085
1.6	1.0	MLE	0.0614	0.0567	0.3953
		MMLE	0.0615	0.0567	0.3969
		MME	0.0672	0.0661	0.4742
		MLE1 (a=1)	0.0616	0.0568	0.3976
2.0	1.0	MLE	0.0643	0.0596	0.4118
		MMLE	0.0643	0.0595	0.4112
		MME	0.0672	0.0659	0.4589
		MLE1 (a=1)	0.0643	0.0596	0.4118
1.4	1.2	MLE	0.0639	0.0587	0.4000
		MMLE	0.0644	0.0591	0.4020
		MME	0.0762	0.0735	0.5330
		MLE1 (a=1)	0.0644	0.0591	0.4021
1.8	1.2	MLE	0.0607	0.0556	0.3195
		MMLE	0.0608	0.0557	0.3923
		MME	0.0677	0.0666	0.4775
		MLE1 (a=1)	0.0605	0.0544	0.3912

Table 3.2.1b : P-value of the Test for a = 1 for Gamma Distribution.

$a \neq 1$	$a = 1$				
	$\lambda_2 = 1.2$ $\sigma^2 = 1.0$	$\lambda_2 = 1.6$ $\sigma^2 = 1.0$	$\lambda_2 = 2.0$ $\sigma^2 = 1.0$	$\lambda_2 = 1.4$ $\sigma^2 = 1.2$	$\lambda_2 = 1.8$ $\sigma^2 = 1.2$
0.0000	0.5059	0.5068	0.5029	0.4755	0.5046

Table 3.2.3 : Estimation of Parameter a for Gamma Distibution.

Method	a				
	0.90	0.95	1.00	1.05	1.10
MLE					
$\lambda_2=1.2$ $\sigma^2=1.0$	0.900032 (5.432E-5)	0.950034 (5.734E-5)	1.000036 (6.035E-5)	1.050038 (6.337E-5)	1.100148 (6.639E-5)
$\lambda_2=1.6$ $\sigma^2=1.0$	0.900043 (3.873E-5)	0.950045 (4.089E-5)	1.000048 (4.304E-5)	1.050050 (4.519E-5)	1.100053 (4.734E-5)
$\lambda_2=2.0$ $\sigma^2=1.0$	0.900025 (3.321E-5)	0.950027 (3.506E-5)	1.000028 (3.690E-5)	1.050030 (3.875E-5)	1.100031 (4.059E-5)
$\lambda_2=1.4$ $\sigma^2=1.2$	0.899956 (4.694E-5)	0.949953 (4.955E-5)	0.999951 (5.216E-5)	1.049948 (5.476E-5)	1.099946 (5.737E-5)
$\lambda_2=1.8$ $\sigma^2=1.2$	0.900048 (3.826E-5)	0.950051 (4.038E-5)	1.000054 (4.251E-5)	1.050056 (4.463E-5)	1.100059 (4.676E-5)
NP					
$\lambda_2=1.2$ $\sigma^2=1.0$	0.899978 (6.048E-5)	0.949977 (6.384E-5)	0.999976 (6.720E-5)	1.049975 (7.056E-5)	1.099974 (7.392E-5)
$\lambda_2=1.6$ $\sigma^2=1.0$	0.900010 (4.298E-5)	0.950011 (4.537E-5)	1.000012 (4.776E-5)	1.050012 (5.014E-5)	1.100013 (5.253E-5)
$\lambda_2=2.0$ $\sigma^2=1.0$	0.900019 (3.504E-5)	0.950020 (3.698E-5)	1.000021 (3.893E-5)	1.050022 (4.088E-5)	1.100023 (4.282E-5)
$\lambda_2=1.4$ $\sigma^2=1.2$	0.899940 (5.462E-5)	0.949936 (5.766E-5)	0.999933 (6.069E-5)	1.049929 (6.372E-5)	1.099926 (6.676E-5)
$\lambda_2=1.8$ $\sigma^2=1.2$	0.900077 (4.200E-5)	0.950081 (4.433E-5)	1.000085 (4.667E-5)	1.050089 (4.900E-5)	1.100094 (5.133E-5)

Table 3.2.4 : Estimation of Parameter λ for Gamma Distribution.

Method	λ				
	1.2 $\begin{pmatrix}\sigma^2=1.00\\ \alpha=1.44\\ \beta=1.20\end{pmatrix}$	1.6 $\begin{pmatrix}\sigma^2=1.00\\ \alpha=2.56\\ \beta=1.60\end{pmatrix}$	2.0 $\begin{pmatrix}\sigma^2=1.00\\ \alpha=4.00\\ \beta=2.00\end{pmatrix}$	1.4 $\begin{pmatrix}\sigma^2=1.20\\ \alpha=1.63\\ \beta=1.17\end{pmatrix}$	1.8 $\begin{pmatrix}\sigma^2=1.20\\ \alpha=2.70\\ \beta=1.50\end{pmatrix}$
MLE	1.208037 (0.0085)	1.608500 (0.0079)	2.006706 (0.0087)	1.397769 (0.0085)	1.811104 (0.0092)
MMLE/MME/NP5	1.202982 (0.0091)	1.604030 (0.0086)	2.005840 (0.0090)	1.398086 (0.0096)	1.818182 (0.0099)
MLE1 / NP6 (a=1.00)	1.199165 (0.0041)	1.598115 (0.0041)	1.998118 (0.0040)	1.401024 (0.0044)	1.797913 (0.0045)
NP1	1.084676 (0.0084)	1.544768 (0.0083)	1.973155 (0.0088)	1.286068 (0.0090)	1.757459 (0.0096)
NP2	1.177647 (0.0089)	1.592482 (0.0085)	1.999070 (0.0090)	1.375283 (0.0095)	1.806048 (0.0099)
NP3	1.223475 (0.0093)	1.614543 (0.0087)	2.012922 (0.0090)	1.416179 (0.0099)	1.830005 (0.0100)
NP4 (a=0.90)	1.209136 (0.0170)	1.588295 (0.0135)	2.012048 (0.0146)	1.391938 (0.0153)	1.834256 (0.0169)
(a=0.95)	1.201494 (0.0130)	1.594062 (0.0106)	2.008016 (0.0111)	1.394457 (0.0126)	1.826248 (0.0134)
(a=1.00)	1.201727 (0.0091)	1.603385 (0.0086)	2.005531 (0.0090)	1.396787 (0.0096)	1.817484 (0.0099)
(a=1.05)	1.200936 (0.0101)	1.604761 (0.0091)	2.009876 (0.0103)	1.391551 (0.0103)	1.818997 (0.0113)
(a=1.10)	1.197906 (0.0126)	1.603085 (0.0117)	2.012701 (0.0131)	1.383108 (0.0132)	1.822590 (0.0142)

Table 3.2.5 : Estimation of Parameter σ^2 using Gamma Distribution.

Method	σ^2				
	1.0 $\begin{pmatrix}\lambda=1.00\\ \alpha=1.44\\ \beta=1.20\end{pmatrix}$	1.0 $\begin{pmatrix}\lambda=1.00\\ \alpha=2.56\\ \beta=1.60\end{pmatrix}$	1.0 $\begin{pmatrix}\lambda=1.00\\ \alpha=4.00\\ \beta=2.00\end{pmatrix}$	1.2 $\begin{pmatrix}\lambda=1.20\\ \alpha=1.63\\ \beta=1.17\end{pmatrix}$	1.2 $\begin{pmatrix}\lambda=1.20\\ \alpha=2.70\\ \beta=1.50\end{pmatrix}$
MLE	1.021608 (0.0148)	1.013028 (0.0112)	1.012103 (0.0107)	1.200248 (0.0162)	1.219498 (0.0137)
MMLE	1.016824 (0.0155)	1.009615 (0.0119)	1.012506 (0.0111)	1.205895 (0.0181)	1.231809 (0.0149)
MME / NP2	1.023243 (0.0160)	1.022263 (0.0124)	1.020728 (0.0116)	1.219549 (0.0188)	1.246526 (0.0157)
MLE1 (a=1.00)	1.002108 (0.0085)	0.999283 (0.0072)	1.003885 (0.0072)	1.201794 (0.0105)	1.201300 (0.0088)
NP1	0.670404 (0.0102)	0.814911 (0.0095)	0.891544 (0.0096)	0.839868 (0.0124)	1.009028 (0.0121)
NP3 (a=1.00)	1.007233 (0.0098)	1.010346 (0.0083)	1.011571 (0.0079)	1.212443 (0.0117)	1.212483 (0.0100)

Table 3.2.6 : Estimation of Parameter α and β for Gamma Distribution.

Method		α,β				
		$\alpha=1.44$ $\beta=1.20$ $\begin{pmatrix}\lambda_2=1.20\\ \sigma^2=1.00\end{pmatrix}$	$\alpha=2.56$ $\beta=1.60$ $\begin{pmatrix}\lambda_2=1.60\\ \sigma^2=1.00\end{pmatrix}$	$\alpha=4.00$ $\beta=2.00$ $\begin{pmatrix}\lambda_2=2.00\\ \sigma^2=1.00\end{pmatrix}$	$\alpha=1.6\dot{3}$ $\beta=1.1\dot{6}$ $\begin{pmatrix}\lambda_2=1.40\\ \sigma^2=1.20\end{pmatrix}$	$\alpha=2.70$ $\beta=1.50$ $\begin{pmatrix}\lambda_2=1.80\\ \sigma^2=1.20\end{pmatrix}$
MLE	α	1.457582 (0.0075)	2.592636 (0.0136)	4.0473212 (0.0236)	1.660543 (0.0092)	2.731305 (0.0147)
	β	1.223120 (0.0102)	1.623430 (0.0116)	2.026710 (0.0151)	1.200915 (0.0098)	1.519065 (0.0110)
MMLE	α	1.455348 (0.0075)	2.590294 (0.0136)	4.040912 (0.0236)	1.658492 (0.0092)	2.729026 (0.0147)
	β	1.229585 (0.0111)	1.628587 (0.0121)	2.027406 (0.0153)	1.202970 (0.0107)	1.513661 (0.0113)
MME	α	1.454651 (0.0105)	2.566237 (0.0164)	4.020009 (0.0267)	1.648634 (0.0118)	2.709396 (0.0180)
	β	1.227785 (0.0122)	1.612746 (0.0131)	2.016829 (0.0165)	1.195216 (0.0118)	1.503147 (0.0129)
MLE1 (a=1.00)	α	1.451199 (0.0075)	2.581415 (0.0135)	4.021688 (0.0232)	1.654360 (0.0092)	2.719091 (0.0146)
	β	1.214268 (0.0074)	1.618508 (0.0095)	2.015269 (0.0123)	1.184694 (0.0078)	1.515193 (0.0090)

Table 3.2.7 : Comparison of Estimates using Parametric and Nonparametric method for Gamma Distribution.

a	$\hat{\lambda}$	DM	MSE	MSEDM	$\hat{\sigma}^2$	DMVAR
0.90	NP1	5	6	3.6	NP1	4
	NP2	2.4	4	3.4	MMLE	2
	NP3	4	4.6	6	MLE	1
	NP4	6	1	1	MME / NP2	3
	MMLE/MME/NP5	2.6	3.2	4.4		
	MLE	1	2.2	2.6		
0.95	NP1	5	6	5	NP1	4
	NP2	2.4	4.2	3.6	MMLE	2
	NP3	4	4.8	5.4	MLE	1
	NP4	6	1	1	MME / NP2	3
	MMLE/MME/NP5	2.6	3	3.8		
	MLE	1	2	2.2		
1.00	NP1	7	6.4	7	NP1	6
	NP2	3.8	4.8	4.2	NP3	2
	NP3	6	4.2	6	MMLE	4
	NP4	4	2	3.8	MLE	3
	MMLE/MME/NP5	4.2	3	4	MME / NP2	5
	MLE1 / NP6	1	6.6	1	MLE1	1
	MLE	2	1	2		
1.05	NP1	5.6	6	5.6	NP1	4
	NP2	2.4	3.6	2.4	MMLE	2
	NP3	4	5	4.2	MLE	1
	NP4	5.4	1	5.2	MME / NP2	3
	MMLE/MME/NP5	2.6	3.4	2.6		
	MLE	1	2	1		
1.10	NP1	5	6	5	NP1	4
	NP2	2.4	3	2.2	MMLE	2
	NP3	4	5	4	MLE	1
	NP4	6	1	6	MME / NP2	3
	MMLE/MME/NP5	2.6	3.6	2.8		
	MLE	1	2.4	1		

Table 3.3.1 : Test for Geometric Process for Weibull Distribution.

α	Is it a HPP ?		Is it a GP ?							
	P_U	S.D.	P_T^U	S.D.	P_D^U	S.D.	P_T^V	S.D.	P_D^V	S.D.
	a≠1									
	.0000	.0000	.4956	.0167	.4965	.0157	.5071	.0164	.4978	.0164
	a=1									
6	.8762	.0047	.4956	.0167	.4965	.0157	.5071	.0164	.4978	.0164
8	.9047	.0036	.4956	.0167	.4965	.0157	.5071	.0164	.4978	.0164
12	.9346	.0025	.4956	.0167	.4965	.0157	.5071	.0164	.4978	.0164
16	.9503	.0019	.4956	.0167	.4965	.0157	.5071	.0164	.4978	.0164

Table 3.3.2 : Test for Weibull Distribution Using EDF Statistic.

Method	α	W^*	U^*	A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D$	$\sqrt{N}V$
MMLE		0.0576	0.0548	0.3902	0.5364	0.5381	0.6120	1.0745
MME	6	0.0601	0.0566	0.4036	0.5372	0.5466	0.6218	1.0838
	8	0.0625	0.0590	0.4220	0.5437	0.5557	0.6278	1.0994
	12	0.0657	0.0624	0.4482	0.5526	0.5667	0.6351	1.1193
	16	0.0678	0.0646	0.4653	0.5581	0.5732	0.6397	1.1313
MLE1 (a=1)		0.0564	0.0536	0.3827	0.5427	0.5347	0.6140	1.0775

Table 3.3.3 : Estimation of Parameter a for Weibull Distribution using Nonparametric (NP).

α	a				
	0.90	0.95	1.00	1.05	1.10
6.0	0.900021 (1.321E-5)	0.950022 (1.394E-5)	1.000023 (1.468E-5)	1.050025 (1.541E-5)	1.100026 (1.614E-5)
8.0	0.900016 (9.906E-6)	0.950017 (1.046E-5)	1.000018 (1.101E-5)	1.050018 (1.156E-5)	1.100019 (1.211E-5)
12.0	0.900011 (6.604E-6)	0.950011 (6.971E-6)	1.000012 (7.338E-6)	1.050012 (7.705E-6)	1.100013 (8.072E-6)
16.0	0.900008 (4.953E-6)	0.950008 (5.228E-6)	1.000009 (5.504E-6)	1.050009 (5.779E-6)	1.100010 (6.054E-6)

Table 3.3.4 : Estimation of Parameter λ for Weibull Distribution.

Method	λ				
	0.154620 (σ^2=9.0E-3, α = 6.0, β = 6.0)	0.092772 (σ^2=3.2E-3, α = 6.0, β = 10.0)	0.235436 (σ^2=1.2E-3, α = 8.0, β = 4.0)	0.159714 (σ^2=2.6E-4, α = 12.0, β = 6.0)	0.120948 (σ^2=8.6E-5, α = 16.0, β = 8.0)
MMLE	0.154949 (2.720E-4)	0.092969 (1.632E-4)	0.235798 (3.117E-4)	0.159872 (1.415E-4)	0.121035 (8.054E-5)
MME/NP5	0.154945 (2.724E-4)	0.092967 (1.635E-4)	0.235791 (3.123E-4)	0.159868 (1.419E-4)	0.121033 (8.079E-5)
MLE1 (a=1.00)	0.154551 (1.216E-4)	0.092730 (7.298E-5)	0.235360 (1.412E-4)	0.159681 (6.503E-5)	0.120929 (3.729E-5)
NP1	0.155085 (2.725E-4)	0.093051 (1.635E-4)	0.235916 (3.122E-4)	0.159902 (1.418E-4)	0.121045 (8.075E-5)
NP2	0.154599 (2.727E-4)	0.092759 (1.636E-4)	0.235543 (3.127E-4)	0.159813 (1.420E-4)	0.121014 (8.084E-5)
NP3	0.155200 (2.724E-4)	0.093120 (1.634E-4)	0.235974 (3.121E-4)	0.159910 (1.418E-4)	0.121047 (8.074E-5)
NP4 (a=0.90)	0.154855 (4.460E-4)	0.092913 (2.676E-4)	0.235686 (5.117E-4)	0.159821 (2.328E-4)	0.121006 (1.327E-4)
(a=0.95)	0.154790 (3.461E-4)	0.092874 (2.076E-4)	0.235615 (3.937E-4)	0.159790 (1.775E-4)	0.120989 (1.007E-4)
(a=1.00)	0.154941 (2.724E-4)	0.092964 (1.634E-4)	0.235788 (3.123E-4)	0.159867 (1.419E-4)	0.121032 (8.078E-5)
(a=1.05)	0.154763 (2.881E-4)	0.092858 (1.728E-4)	0.235591 (3.360E-4)	0.159779 (1.556E-4)	0.120983 (8.954E-5)
(a=1.10)	0.154564 (3.741E-4)	0.092738 (2.244E-4)	0.235362 (4.364E-4)	0.159676 (2.022E-4)	0.120924 (1.164E-4)
NP6 (a=1.00)	0.154539 (1.224E-4)	0.092724 (7.346E-5)	0.235341 (1.424E-4)	0.159670 (6.576E-5)	0.120922 (3.777E-5)

Table 3.3.5 : Estimation of Parameter σ^2 for Weibull Distribution.

Method	σ^2				
	8.977764E-4 (λ=0.1543, α= 6.0, β= 6.0)	3.231635E-4 (λ=0.0928, α= 6.0, β= 10.0)	1.220198E-3 (λ=0.2354, α= 8.0, β= 4.0)	2.613301E-4 (λ=0.1597, α= 12.0, β= 6.0)	8.642989E-5 (λ=0.1209, α= 16.0, β= 8.0)
MMLE	8.942002E-4 (4.975E-6)	3.219121E-4 (1.791E-6)	1.213867E-3 (6.448E-6)	2.596914E-4 (1.363E-6)	8.585047E-5 (4.540E-7)
MME/NP2	9.001971E-4 (5.566E-6)	3.240710E-4 (2.004E-6)	1.222571E-3 (7.659E-6)	2.617067E-4 (1.740E-6)	8.654635E-5 (6.030E-7)
MLE1 (a=1.00)	8.890244E-4 (4.162E-4)	3.200488E-4 (1.498E-6)	1.208845E-3 (5.801E-6)	2.589979E-4 (1.294E-5)	8.56767.E-5 (4.389E-7)
NP1	1.057326E-3 (8.764E-6)	3.806375E-4 (3.155E-6)	1.402613E-3 (1.150E-5)	2.904705E-4 (2.386E-6)	9.411336E-5 (7.771E-7)
NP3 (a=1.00)	8.973377E-4 (4.885E-6)	3.230416E-4 (1.758E-6)	1.221346E-3 (7.183E-6)	2.619806E-4 (1.703E-6)	8.672174E-5 (5.979E-7)

Table 3.3.6 : Estimation of Parameter α and β for Weibull Distribution.

Meth		α,β				
		α= 6.0, β= 6.0 (λ_2=0.1546, σ^2=8.9E-4)	α= 6.0, β=10.0 (λ_2=0.0928, σ^2=3.2E-4)	α= 8.0, β= 4.0 (λ_2=0.2354, σ^2=1.2E-3)	α=12.0, β= 6.0 (λ_2=0.1597, σ^2=2.6E-4)	α=16.0, β= 8.0 (λ_2=0.1209, σ^2=8.6E-5)
MMLE	α	6.045409 (0.0174)	6.045409 (0.0174)	8.060545 (0.0232)	12.090817 (0.0348)	16.120863 (0.0463)
	β	5.994613 (0.0103)	9.991022 (0.0172)	3.996976 (0.0051)	5.996645 (0.0051)	7.996481 (0.0052)
MME	α	6.028912 (0.0196)	6.028912 (0.0196)	8.040396 (0.0274)	12.065125 (0.0434)	16.091092 (0.0597)
	β	5.993726 (0.0103)	9.989544 (0.0172)	3.996468 (0.0052)	5.996073 (0.0052)	7.995878 (0.0052)
MLE1 (a=1.00)	α	6.045594 (0.0171)	6.045594 (0.0171)	8.060792 (0.0228)	12.091188 (0.0342)	16.121584 (0.0456)
	β	6.005706 (0.0044)	10.009511 (0.0073)	4.002793 (0.0022)	6.002733 (0.0022)	8.002703 (0.0022)

Table 3.3.7 : Comparison of Estimates using Parametric and Nonparametric method for Weibull Distribution.

a	$\hat{\lambda}$	DM	MSE	MSEDM	$\hat{\sigma}^2$	DMVAR
0.90	NP1	3.6	3.6	3.6	NP1	3
	NP2	4.2	6	6	MMLE	1
	NP3	3.8	2.6	2.4	MME / NP2	2
	NP4	6	1	1		
	MMLE	1	2.8	3		
	MME / NP5	2.4	5	5		
0.95	NP1	3.6	3.8	3.6	NP1	3
	NP2	4.2	6	6	MMLE	1
	NP3	3.8	4.4	4.2	MME / NP2	2
	NP4	6	1	1		
	MMLE	1	2	2		
	MME / NP5	2.4	3.8	4.2		
1.00	NP1	6.4	3.6	6.4	NP1	5
	NP2	7.2	5.6	7.2	NP3	3
	NP3	6.6	4.6	6.6	MMLE	2
	NP4	4.4	1	4.4	MME / NP2	4
	NP6	2	7	2	MLE1	1
	MMLE	3	4.2	3		
	MME / NP5	5.4	2	5.4		
	MLE1	1	8	1		
1.05	NP1	3.6	4.8	3.8	NP1	3
	NP2	4.2	2	3.8	MMLE	1
	NP3	3.8	5.8	4.4	MME / NP2	2
	NP4	6	1	5.6		
	MMLE	1	4	1		
	MME / NP5	2.4	3.4	2.4		
1.10	NP1	3.6	4.8	3.8	NP1	3
	NP2	4.2	2	3.8	MMLE	1
	NP3	3.8	5.8	4	MME	2
	NP4	6	1	6		
	MMLE	1	4	1		
	MME / NP5	2.4	3.4	2.4		

Table 3.4.1 : Test for Geometric Process for Lognormal Distribution.

τ^2	Is it a HPP ?		Is it a GP ?							
	P_U	S.D.	P_T^U	S.D.	P_D^U	S.D.	P_T^V	S.D.	P_D^V	S.D.
	a≠1									
	.0000	.0000	.5094	.0170	.5085	.0164	.4885	.0173	.4630	.0166
	a=1									
0.10	.7977	.0083	.5094	.0170	.5085	.0164	.4885	.0173	.4630	.0166
0.20	.7159	.0111	.5094	.0170	.5085	.0164	.4885	.0173	.4630	.0166
0.40	.6067	.0144	.5094	.0170	.5085	.0164	.4885	.0173	.4630	.0166
0.60	.5288	.0160	.5094	.0170	.5085	.0164	.4885	.0173	.4630	.0166

Table 3.4.2 : Test for Lognormal Distribution Using EDF Statistic.

Method	W^*	U^*	A^*	D^*	V^*
MLE / MMLE /MME	0.0604	0.0568	0.3914	0.6282	1.0895
MLE1 (a=1.00)	0.0609	0.0574	0.3927	0.6370	1.1033

Table 3.4.3 : Estimation of Parameter a for Lognormal Distibution using both Nonparametric (NP) and Parametric (MLE) method.

τ^2	a				
	0.90	0.95	1.00	1.05	1.10
0.1	0.900033 (1.982E-5)	0.950035 (2.0922-5)	1.000037 (2.202E-5)	1.050039 (2.312E-5)	1.100041 (2.422E-5)
0.4	0.900067 (3.964E-5)	0.950070 (4.184E-5)	1.000074 (4.404E-5)	1.050078 (4.625E-5)	1.100081 (4.850E-5)
0.6	0.900082 (4.855E-5)	0.950086 (5.125E-5)	1.000091 (5.394E-5)	1.050095 (5.664E-5)	1.100100 (5.934E-5)
1.2	0.900047 (2.803E-5)	0.950050 (2.959E-5)	1.000052 (3.114E-5)	1.050055 (3.270E-5)	1.100057 (3.426E-5)

Table 3.4.4 : Estimation of Parameter λ for Lognormal Distribution.

Method	λ				
	1.822119	1.733253	2.225541	3.004166	9.974182
	$\begin{pmatrix}\sigma^2=1.633\\ \mu_2=0.4\\ \tau^2=0.4\end{pmatrix}$	$\begin{pmatrix}\sigma^2=0.316\\ \mu_2=0.5\\ \tau^2=0.1\end{pmatrix}$	$\begin{pmatrix}\sigma^2=4.072\\ \mu_2=0.5\\ \tau^2=0.6\end{pmatrix}$	$\begin{pmatrix}\sigma^2=1.998\\ \mu_2=1.0\\ \tau^2=0.2\end{pmatrix}$	$\begin{pmatrix}\sigma^2=81.79\\ \mu_2=2.0\\ \tau^2=0.6\end{pmatrix}$
MLE/MMLE	1.839439 (0.0095)	1.739762 (0.0044)	2.254183 (0.0144)	3.021809 (0.0108)	10.102546 (0.0646)
MME/NP5	1.839022 (0.0095)	1.739729 (0.0044)	2.253099 (0.0146)	3.021612 (0.0109)	10.097688 (0.0654)
MLE1 (a=1.00)	1.822842 (0.0053)	1.732633 (0.0024)	2.228248 (0.0083)	3.003563 (0.0059)	9.863173 (0.0371)
NP1	1.748715 (0.0088)	1.732882 (0.0043)	2.042937 (0.0126)	2.977051 (0.0106)	9.155799 (0.0565)
NP2	1.895090 (0.0103)	1.744100 (0.0044)	2.388796 (0.0172)	3.048554 (0.0111)	10.705814 (0.0770)
NP3	1.809897 (0.0092)	1.738470 (0.0044)	2.175033 (0.0137)	3.010280 (0.0108)	9.747822 (0.0612)
NP4 (a=0.90)	1.832881 (0.0181)	1.738098 (0.0079)	2.241058 (0.0285)	3.016561 (0.0200)	10.043724 (0.1277)
(a=0.95)	1.830015 (0.0129)	1.736676 (0.0057)	2.578870 (0.0137)	3.013069 (0.0143)	10.026126 (0.0908)
(a=1.00)	1.839131 (0.0095)	1.739738 (0.0044)	2.253377 (0.0146)	3.021664 (0.0109)	10.098937 (0.0656)
(a=1.05)	1.826821 (0.0123)	1.734657 (0.0055)	2.233393 (0.0191)	3.008540 (0.0137)	10.009374 (0.0856)
(a=1.10)	1.819165 (0.0160)	1.731831 (0.0072)	2.220363 (0.0247)	3.000886 (0.0179)	9.950975 (0.1107)
NP6	1.822511 (0.0054)	1.732606 (0.0024)	2.227402 (0.0086)	3.003406 (0.0060)	9.982523 (0.0386)

Table 3.4.5 : Estimation of Parameter σ^2 for Lognormal Distribution.

Method	σ^2				
	1.632916 $\begin{pmatrix}\lambda=1.822\\ \mu_2=0.4\\ \tau^2=0.4\end{pmatrix}$	0.315951 $\begin{pmatrix}\lambda=1.733\\ \mu_2=0.5\\ \tau^2=0.1\end{pmatrix}$	4.071981 $\begin{pmatrix}\lambda=2.226\\ \mu_2=0.5\\ \tau^2=0.6\end{pmatrix}$	1.998163 $\begin{pmatrix}\lambda=3.004\\ \mu_2=1.0\\ \tau^2=0.2\end{pmatrix}$	81.787926 $\begin{pmatrix}\lambda=9.974\\ \mu_2=2.0\\ \tau^2=0.6\end{pmatrix}$
MLE/MMLE	1.688929 (0.0240)	0.319182 (0.0027)	4.279571 (0.0744)	2.034853 (0.0217)	85.957484 (1.4936)
MME/NP2	1.681476 (0.0296)	0.320116 (0.0031)	4.227434 (0.0969)	2.036691 (0.0252)	84.910282 (1.9470)
MLE1 (a=1.00)	1.659288 (0.0186)	0.317760 (0.0024)	4.174426 (0.0567)	2.015635 (0.0176)	83.845597 (1.1381)
NP1	0.922414 (0.0109)	0.276923 (0.0022)	1.707000 (0.0236)	1.514759 (0.0142)	34.286010 (0.4745)
NP3 (a=1.00)	1.655583 (0.0257)	0.318825 (0.0027)	4.139541 (0.0854)	2.019283 (0.0219)	83.144905 (1.7147)

Table 3.4.6 : Estimation of Parameter μ and τ^2 for Lognormal Distribution.

Method		μ, τ^2				
		$\mu=0.40$ $\tau^2=0.40$ $\begin{pmatrix}\lambda_2=1.822\\ \sigma^2=1.633\end{pmatrix}$	$\mu=0.50$ $\tau^2=0.10$ $\begin{pmatrix}\lambda_2=1.733\\ \sigma^2=0.316\end{pmatrix}$	$\mu=0.50$ $\tau^2=0.60$ $\begin{pmatrix}\lambda_2=2.226\\ \sigma^2=4.072\end{pmatrix}$	$\mu=1.00$ $\tau^2=0.20$ $\begin{pmatrix}\lambda_2=3.004\\ \sigma^2=1.998\end{pmatrix}$	$\mu=2.00$ $\tau^2=0.60$ $\begin{pmatrix}\lambda_2=9.974\\ \sigma^2=81.79\end{pmatrix}$
MLE/ MMLE	μ	0.405695 (0.0050)	0.502847 (0.0025)	0.506974 (0.0061)	1.004027 (0.0035)	2.006974 (0.0061)
	τ^2	0.399632 (0.0024)	0.099908 (0.0006)	0.599448 (0.0036)	0.199816 (0.0012)	0.599448 (0.0036)
MME	μ	0.407117 (0.0051)	0.502708 (0.0025)	0.512391 (0.0065)	1.004002 (0.0036)	2.012391 (0.0065)
	τ^2	0.396223 (0.0038)	0.100143 (0.0071)	0.587351 (0.0065)	0.199715 (0.0016)	0.587351 (0.0065)
MLE1 (a=1.00)	μ	0.398332 (0.0027)	0.499166 (0.0013)	0.497957 (0.0032)	0.998820 (0.0019)	1.997957 (0.0032)
	τ^2	0.401602 (0.0024)	0.100401 (0.0006)	0.602404 (0.0036)	0.200801 (0.0012)	0.602404 (0.0036)

Table 3.4.7 : Comparison of Estimates using Parametric and Nonparametric method for Lognormal Distribution.

a	$\hat{\lambda}$	DM	MSE	MSEDM	$\hat{\sigma}^2$	DMVAR
0.90	NP1	3	6	2	NP1	3
	NP2	5	4.2	6	MME / NP2	2
	NP3	1.2	4.4	3	MLE / MMLE	1
	NP4	6	1	1		
	MME / NP5	3.4	2.2	4		
	MLE / MMLE	2.4	3.2	5		
0.95	NP1	3	6	3.6	NP1	3
	NP2	5.4	5	6	MME / NP2	2
	NP3	1.2	4	2.2	MLE / MMLE	1
	NP4	5.6	1	1		
	MME / NP5	3.4	2	3.8		
	MLE / MMLE	2.4	3	4.4		
1.00	NP1	5.6	7.2	5.8	NP1	5
	NP2	8	5.8	8	MME / NP2	4
	NP3	3.2	4	3.2	NP3	3
	NP4	6.4	1	6.4	MLE / MMLE	2
	MME / NP5	5.4	2	5.4	MLE1	1
	NP6	2	6	2		
	MLE1	1	7	1		
	MLE / MMLE	4.4	3	4.2		
1.05	NP1	3	4.4	3	NP1	3
	NP2	5	6	5.4	MME / NP2	2
	NP3	1.2	3	1.2	MLE / MMLE	1
	NP4	6	1	5.6		
	MME / NP5	3.4	2.8	3.4		
	MLE / MMLE	2.4	3.8	2.4		
1.10	NP1	3	3.8	3	NP1	3
	NP2	5	6	5	MME / NP2	2
	NP3	1.2	2.4	1.2	MLE / MMLE	1
	NP4	6	1	6		
	MME / NP5	3.4	4	3.4		
	MLE / MMLE	2.4	3.8	2.4		

Table 3.5.1 : Results of Simulation Studies for Estimators $\hat{\lambda}$.

Dist.	DM			MSE			MSEDM		
	a<1	a>1	a=1	a<1	a>1	a=1	a<1	a>1	a=1
Exp.	MLE	MLE	MLE1	*NP4*	*NP4*	MLE	*NP4*	MLE	MLE1
Gam.	MLE	MLE	MLE1	*NP4*	*NP4*	MLE	*NP4*	MLE	MLE1
Wei.	MMLE	MMLE	MLE1	*NP4*	*NP4*	*NP4*	*NP4*	MMLE	MLE1
Log.	*NP3*	*NP3*	MLE1	*NP4*	*NP4*	*NP4*	*NP4*	*NP3*	MLE1

Table 4.1.1 : Test for Geometric Process for Coal data.

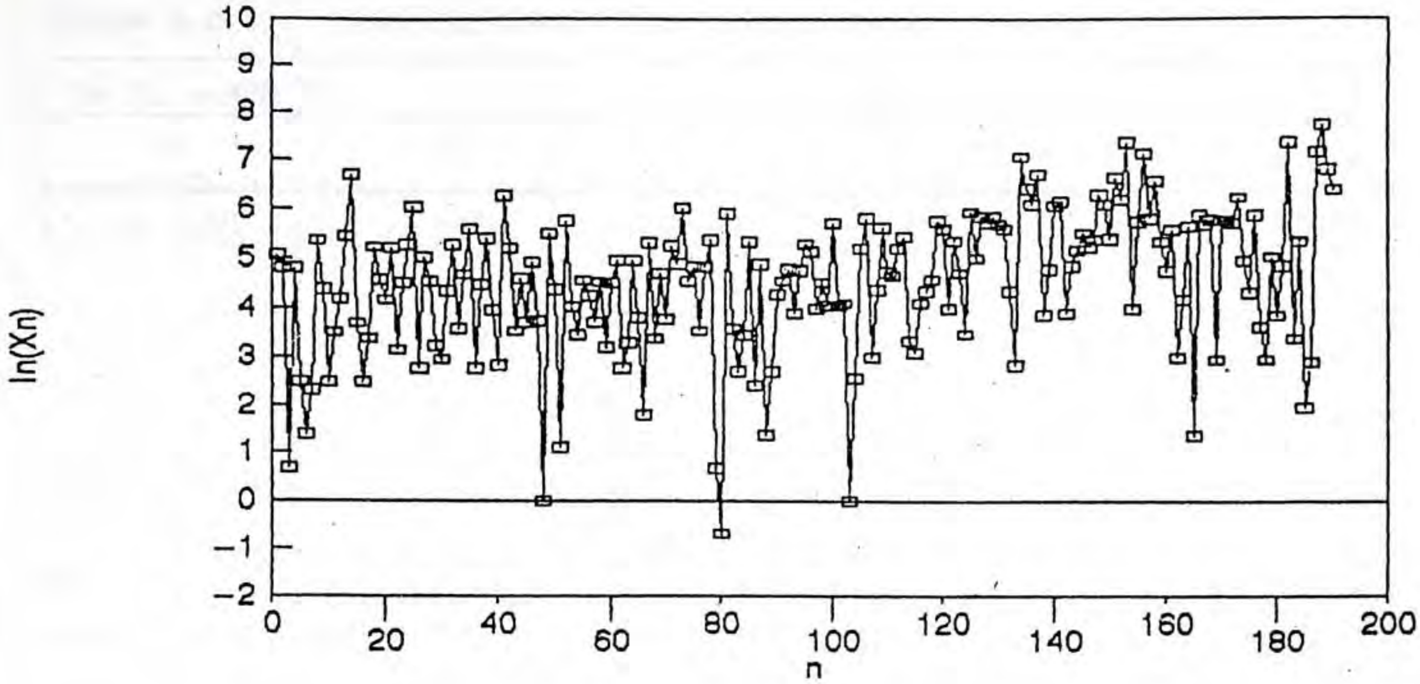
Is it a HPP ?	Is it a GP ?				Is a=1 ?
P_u	P_T^U	P_D^U	P_T^V	P_D^V	P_t
0.0000	0.4611	0.4795	0.1404	0.4795	0.0000

Table 4.1.2 : Test for Distribution for Coal data.

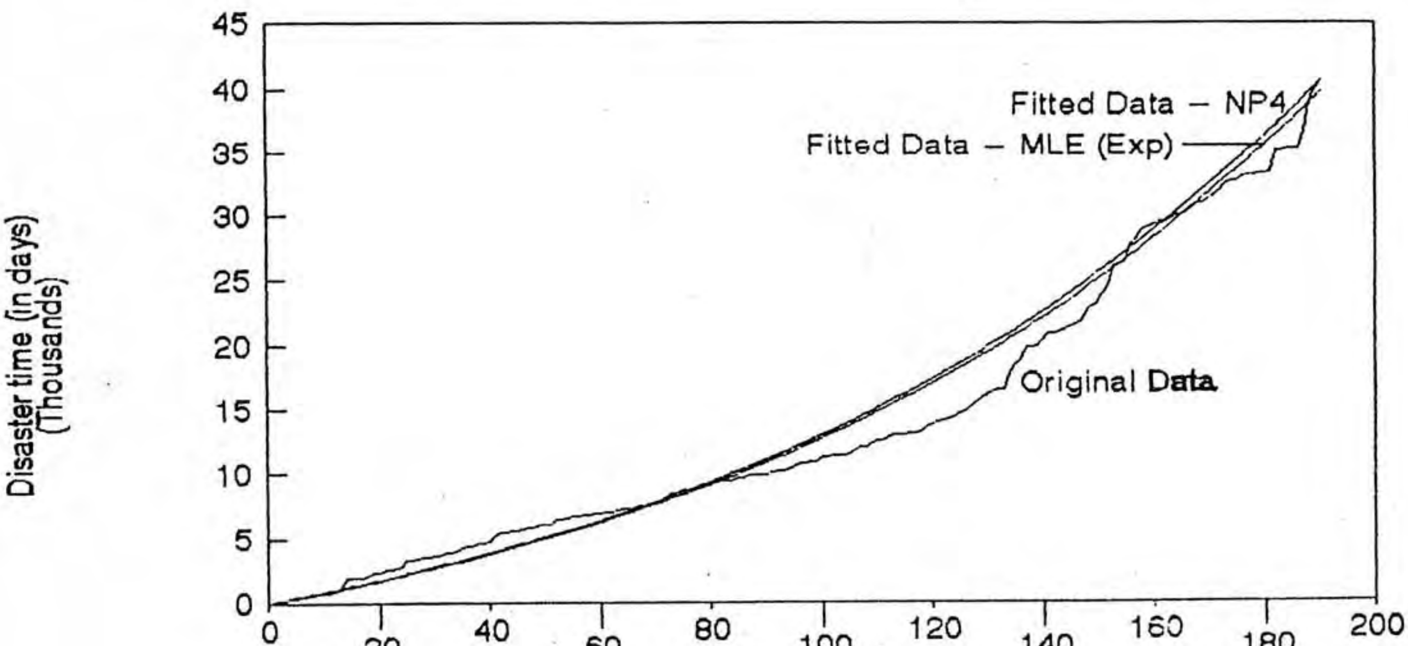
Method	KS-Test							F-Test	χ^2 GOF-Test (DF)
	W/W^*	U/U^*	A/A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D/D^*$	$\sqrt{N}V/V^*$		
MLE_E	*.1503	*.0616	*.9933			*.8558	⊙1.6486	.4506	.5123(18)
$MMLE_E$	*.1503	*.0616	*.9933			*.8557	⊙1.6484	.4508	.5123(18)
$MLE1_E$	.8120	.2304	4.8619			1.4499	⊙2.8343	.0003	.0950(18)
MLE_G	*.0420	*.0371	*.2776						.8538(17)
$MMLE_G$	*.0420	*.0371	*.2776						.8538(17)
MME_G	.1903	⊙.1350	1.2320						.4723(17)
$MLE1_G$	⊙.1441	*.0881	.9917						.2689(17)
$MMLE_W$	*.0403	*.0403	*.2479	*.5627	*.4948	*.5627	*1.0574		.7256(17)
MME_W	*.1025	*.0876	*.5929	⊙.7915	*.4091	*.7915	*1.2007		.4579(17)
$MLE1_W$	*.0577	*.0533	*.5156	*.5472	*.6392	*.6392	*1.1864		.3255(17)
MLE_L	.4169	.3223	2.4824			1.4222	2.2242		.0175(17)
MME_L	.4169	.3223	2.4824			1.4222	2.2242		.0175(17)
$MLE1_L$	.2321	.1783	1.4528			1.0261	1.6652		.3888(17)

Table 4.1.3 : Estimation of Parameters for Coal data.

Method	$\hat{a}$	$\hat{\lambda}$				$\hat{\sigma}^2$	$\hat{\alpha}$ or $\hat{\mu}$	$\hat{\beta}$ or $\hat{\tau}^2$
		Value	MSE	Rank	MSER			
NP1	0.990913	59.541	86591	14	6.05%	2983		
NP2		78.788	81719	2	0.08%	8700		
NP3		77.246	81867	12	0.26%			
NP4		79.696	81652	1	-			
NP5		78.009	81789	3	0.17%			
NP6	1.000000	213.418	97794	15	19.8%	98311		
MLE_E	0.990914	78.017	81790	9	0.17%	6087		0.0128177
$MMLE_E$	0.990913	78.009	81789	3	0.17%	6085		0.0128190
$MLE1_E$	1.000000	213.418	97794	15	19.8%	45547		0.0046856
MLE_G	0.990914	78.017	81790	9	0.17%	7144	0.85203	0.0109210
$MMLE_G$	0.990913	78.009	81789	3	0.17%	7142	0.85203	0.0109222
MME_G	0.990913	78.009	81789	3	0.17%	8700	0.69940	0.0089657
$MLE1_G$	1.000000	213.481	97794	15	19.8%	63163	0.72110	0.0033788
$MMLE_W$	0.990913	77.883	81801	11	0.18%	7626	0.89358	0.0135632
MME_W	0.990913	78.009	81789	3	0.17%	8700	0.84043	0.0140468
$MLE1_W$	1.000000	210.791	97801	18	19.8%	72000	0.79277	0.0054104
MLE_L	0.990913	102.682	84830	13	3.89%	62198	3.66593	1.9314042
MME_L	0.990913	78.009	81789	3	0.17%	8700	3.91292	0.8878047
$MLE1_L$	1.000000	275.797	101685	19	24.5%	598239	4.52861	2.1821033



Graph 4.1.1 : Plot of $\ln X_n$ against n for Coal mining disasters data.



Graph 4.1.2 : Plot of Coal mining disasters data and their fitted values using NP4 and MLE (Exp)

Table 4.2.1 : Test for Geometric Process for Air1 data.

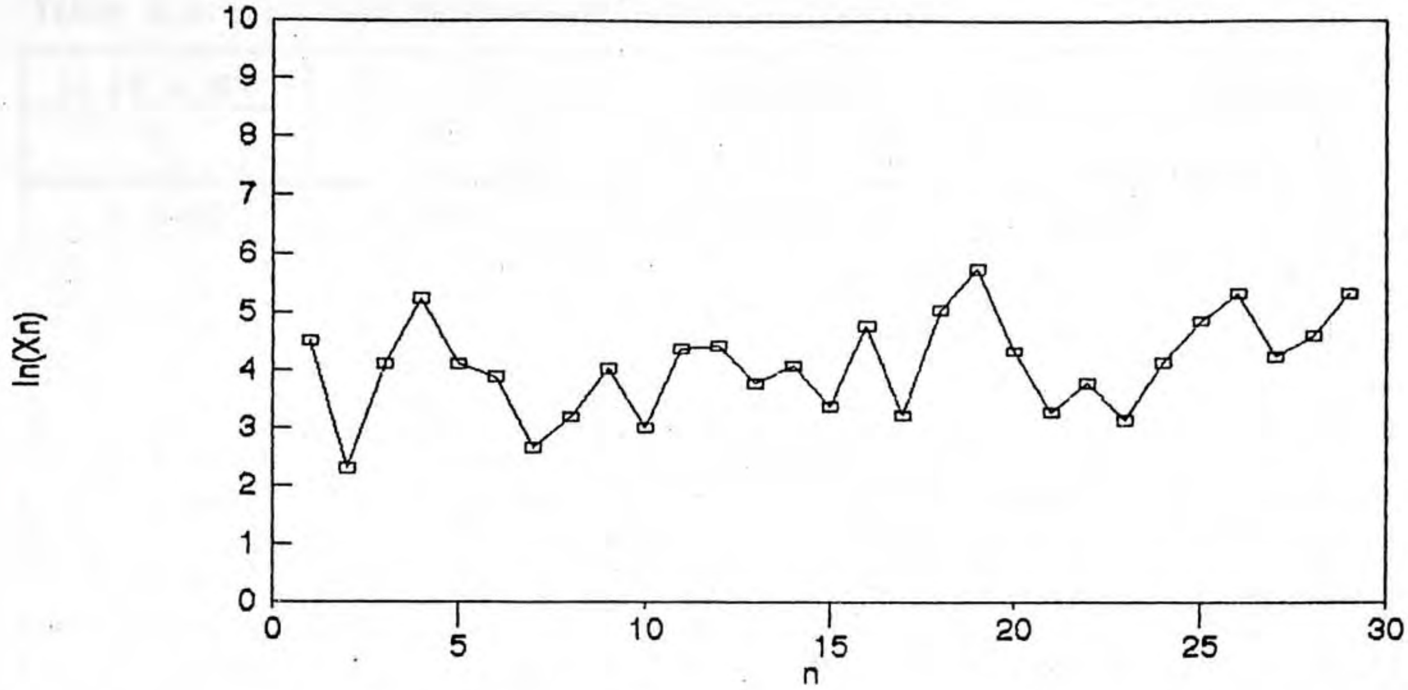
Is it a HPP ?	Is it a GP ?				Is a=1 ?
P_u	P_T^U	P_D^U	P_T^V	P_D^V	P_t
0.1388	0.4969	0.1797	0.1742	0.6547	0.0599

Table 4.2.2 : Test for Distribution for Air1 data.

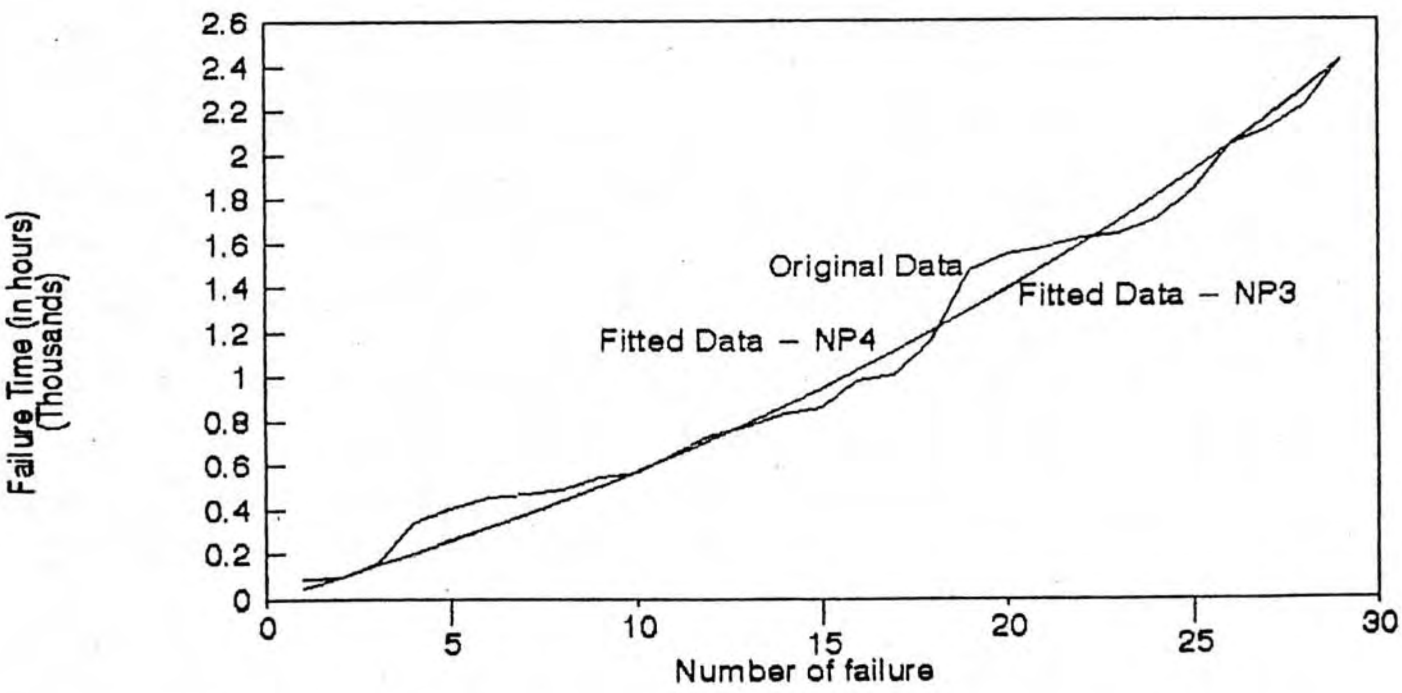
Method	KS-Test							F-Test	χ^2 GOF-Test (DF)
	W/W^*	U/U^*	A/A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D/D^*$	$\sqrt{N}V/V^*$		
MLE_E	*.1549	*.1047	⊙1.0728			*.9146	⊙1.6570	.1512	.3098 (3)
$MMLE_E$	*.1560	*.1078	⊙1.0955			*.9807	1.7882	.1651	.2690 (3)
$MLE1_E$	*.1223	*.0898	*0.8268			*.7866	*1.4032	.3417	.1296 (3)
MLE_G	*.0376	*.0320	*.3319						.4683 (2)
$MMLE_G$	*.0432	*.0367	*.3783						.4683 (2)
MME_G	*.0380	*.0378	*.3697						.9334 (2)
$MLE1_G$	*.0484	*.0415	*.3141						.1664 (2)
$MMLE_W$	*.0477	*.0424	*.4423	*.4977	*.5479	*.5479	*1.0456		.4828 (2)
MME_W	*.0436	*.0440	*.4736	*.6304	*.4933	*.6304	*1.1237		.7855 (2)
$MLE1_W$	*.0583	*.0526	*.3791	*.3373	*.5212	*.5212	*.8585		.8012 (3)
MLE_L	*.0468	*.0467	*.3228			*.5212	*.9969		.4683 (2)
MME_L	*.0468	*.0467	*.3228			*.5212	*.9969		.4683 (2)
$MLE1_L$	*.0336	*.0332	*.2089			*.4823	*.9357		.1664 (2)

Table 4.2.3 : Estimation of Parameters for Air1 data.

Method	$\hat{a}$	$\hat{\lambda}$ Value	MSE	Rank	MSER	$\hat{\sigma}^2$	$\hat{\alpha}$ or $\hat{\mu}$	$\hat{\beta}$ or $\hat{\tau}^2$
NP1	0.965282	45.4346	4300.60	14	0.79%	883.7		
NP2		50.4993	4277.45	11	0.25%	1630.7		
NP3		48.7260	4266.97	1	-			
NP4		48.7667	4266.98	2	2E-6			
NP5		49.3012	4268.17	3	3E-5			
NP6	1.000000	83.5172	4840.59	15	13.4%	5013.5		
MLE_E	0.970870	53.3810	4286.96	12	0.47%	2849.5		0.0187333
$MMLE_E$	0.965282	49.3012	4268.17	3	3E-5	2340.6		0.0208350
$MLE1_E$	1.000000	83.5172	4840.59	15	13.4%	6975.1		0.0119736
MLE_G	0.970870	53.3810	4286.96	12	0.47%	1542.4	1.84749	0.0346094
$MMLE_G$	0.965282	49.3012	4268.17	3	3E-5	1321.0	1.83994	0.0373205
MME_G	0.965282	49.3012	4268.17	3	3E-5	1630.7	1.49050	0.0302326
$MLE1_G$	1.000000	83.5172	4840.59	15	13.4%	4174.3	1.67099	0.0200077
$MMLE_W$	0.965282	49.6671	4270.04	9	0.07%	1378.0	1.35255	0.0184565
MME_W	0.965282	49.3012	4268.17	3	3E-5	1630.7	1.22737	0.0189728
$MLE1_W$	1.000000	84.0471	4840.88	18	13.5%	4292.2	1.29331	0.0110003
MLE_L	0.965282	49.7988	4270.92	10	0.09%	2091.1	3.60223	0.6115204
MME_L	0.965282	49.3012	4268.17	3	3E-5	1630.7	3.64126	0.5132717
$MLE1_L$	1.000000	85.3170	4843.83	19	13.5%	7363.2	4.09692	0.6989175



Graph 4.2.1 : Plot of $\ln X_n$ against n for Air1 data.



Graph 4.2.2 : Plot of Air1 data and their fitted values using NP3 and NP4.

Table 4.3.1 : Test for Geometric Process for No3 data.

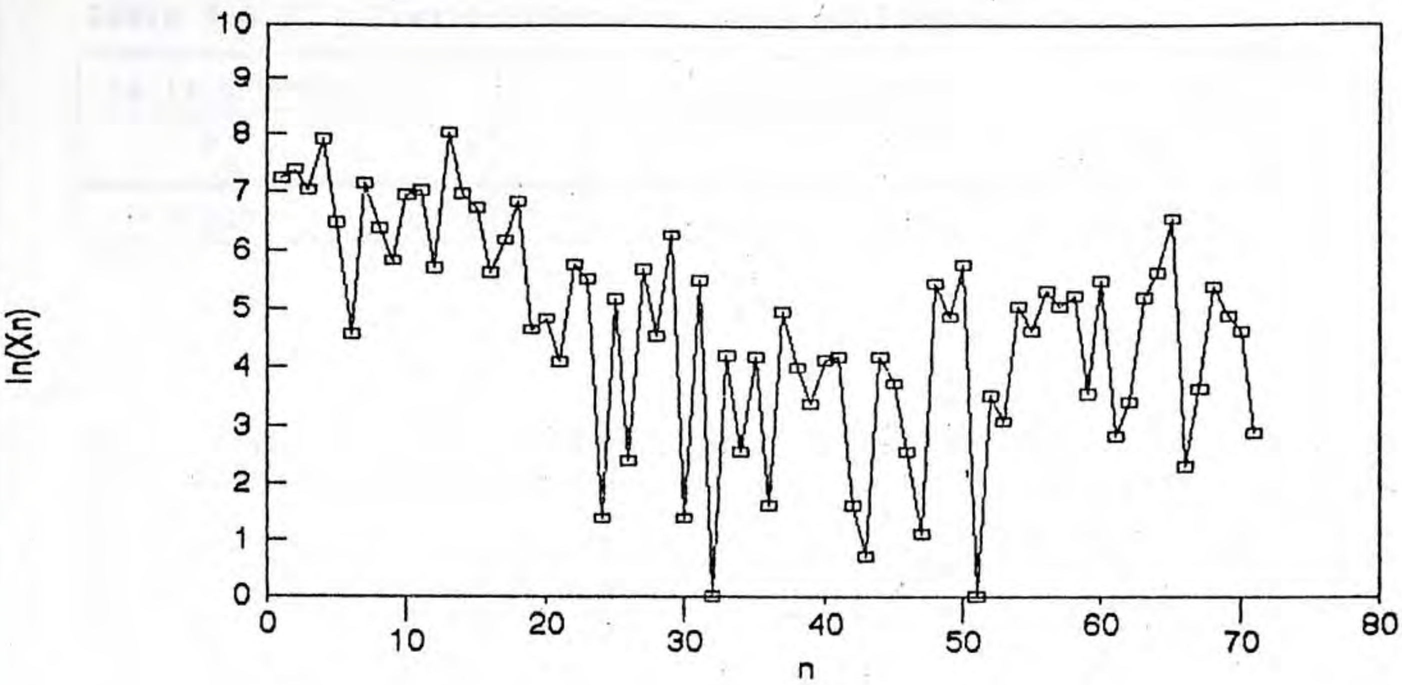
Is it a HPP ?	Is it a GP ?				Is a=1 ?
P_u	P_T^U	P_D^U	P_T^V	P_D^V	P_t
0.0000	0.2168	0.5637	0.6806	0.5637	0.0001

Table 4.3.2 : Test for Distribution for No3 data.

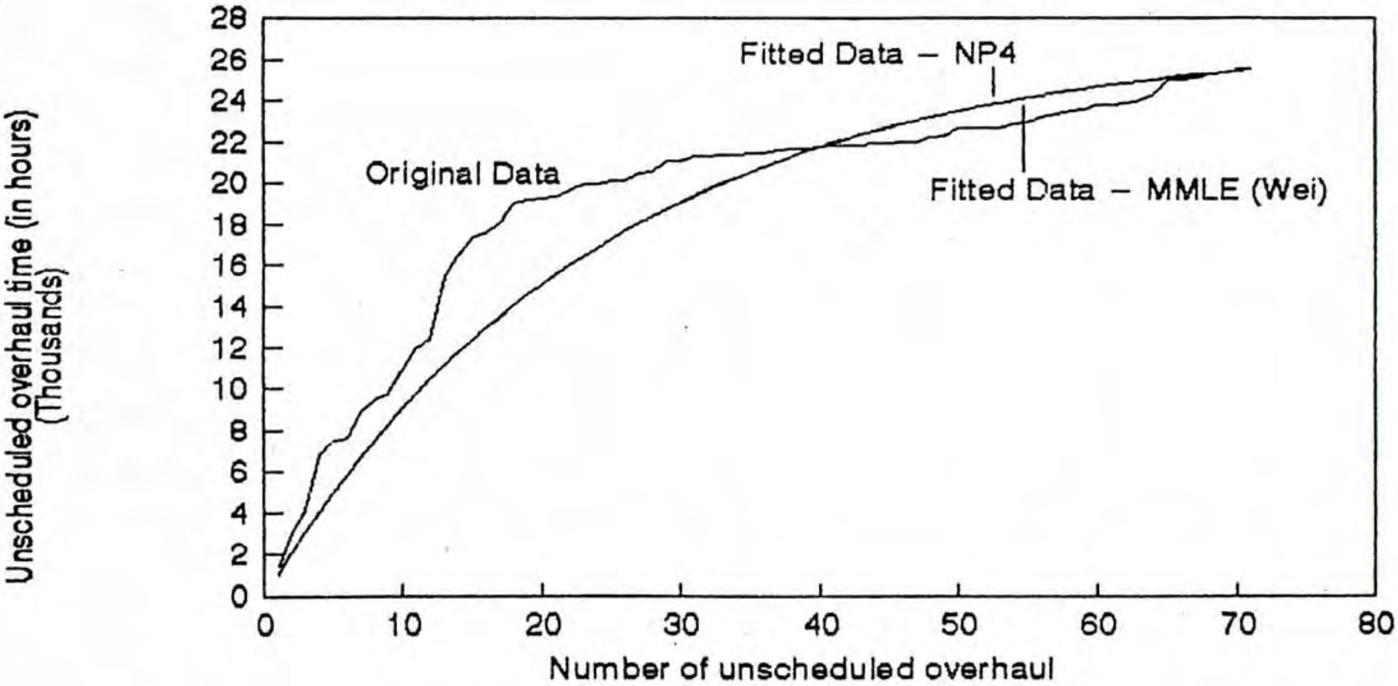
Method	KS-Test							F-Test	χ^2 GOF-Test (DF)
	W/W^*	U/U^*	A/A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D/D^*$	$\sqrt{N}V/V^*$		
MLE_E	.3092	.1964	2.4990			1.3667	2.6235	.0250	.0502 (7)
$MMLE_E$	.3110	⊙.1690	2.4577			1.1832	2.2581	.0077	.2923 (7)
$MLE1_E$	1.3295	.5105	9.2546			1.9168	3.7188	.0000	.0000 (4)
MLE_G	*.0652	*.0651	*.4598						.3072 (7)
$MMLE_G$	*.0500	*.0497	*.3664						.3248 (6)
MME_G	⊙.1095	*.0783	⊙.6905						.2427 (7)
$MLE1_G$	*.1095	*.0760	*.6099						.4407 (2)
$MMLE_W$	*.0605	*.0594	*.4483	*.6581	*.4700	*.6581	*1.1282		.0930 (7)
MME_W	*.0671	*.0670	*.4641	*.6764	*.4339	*.6764	*1.1104		.3835 (7)
$MLE1_W$	*.0331	*.0307	*.2421	*.3561	*.5888	*.5888	*.9449		.6538 (6)
MLE_L	.2634	.2119	1.7542			1.0258	1.6770		.0023 (5)
MME_L	.2634	.2119	1.7542			1.0258	1.6770		.0023 (5)
$MLE1_L$	*.0997	*.0812	⊙.6414			*.7364	*1.2343		.7384 (7)

Table 4.2.3 : Estimation of Parameters for No3 data.

Method	$\hat{a}$	$\hat{\lambda}$				$\hat{\sigma}^2$	$\hat{\alpha}$ or $\hat{\mu}$	$\hat{\beta}$ or $\hat{\tau}^2$
		Value	MSE	Rank	MSER			
NP1	1.041649	721.09	237790	13	21.3%	538906		
NP2		1073.31	196336	9	0.18%	2110312		
NP3		1056.88	197282	10	0.66%			
NP4		1079.90	195984	1	-			
NP5		1076.21	196179	3	0.10%			
NP6	1.000000	359.41	332153	15	69.5%	336898		
MLE_E	1.035476	864.57	215623	11	10.0%	747481		0.0011566
$MMLE_E$	1.041649	1076.21	196179	3	0.10%	1158228		0.0009292
$MLE1_E$	1.000000	359.41	332153	15	69.5%	129174		0.0027823
MLE_G	1.035476	864.57	215623	11	10.0%	1115791	0.66991	0.0007748
$MMLE_G$	1.041649	1076.21	196179	3	0.10%	1747129	0.66293	0.0006160
MME_G	1.041649	1076.21	196179	3	0.10%	2110312	0.54884	0.0005100
$MLE1_G$	1.000000	359.41	332153	15	69.5%	254776	0.50701	0.0014107
$MMLE_W$	1.041649	1079.40	196010	2	0.01%	2038315	0.76546	0.0010851
MME_W	1.041649	1076.21	196179	3	0.10%	2110312	0.75152	0.0011045
$MLE1_W$	1.000000	355.13	332171	18	69.5%	345349	0.63032	0.0039852
MLE_L	1.041649	1776.29	246508	14	25.8%	50818997	6.06256	2.8394526
MME_L	1.041649	1076.21	196179	3	0.10%	2110312	6.46248	1.0374518
$MLE1_L$	1.000000	604.11	392033	19	100%	12198550	4.63437	3.5387878



Graph 4.3.1 : Plot of $\ln X_n$ against n for No3 data.



Graph 4.3.2 : Plot of No3 data and their fitted values using NP4 and MMLE (Wei).

Table 4.4.1 : Test for Geometric Process for Air2 data.

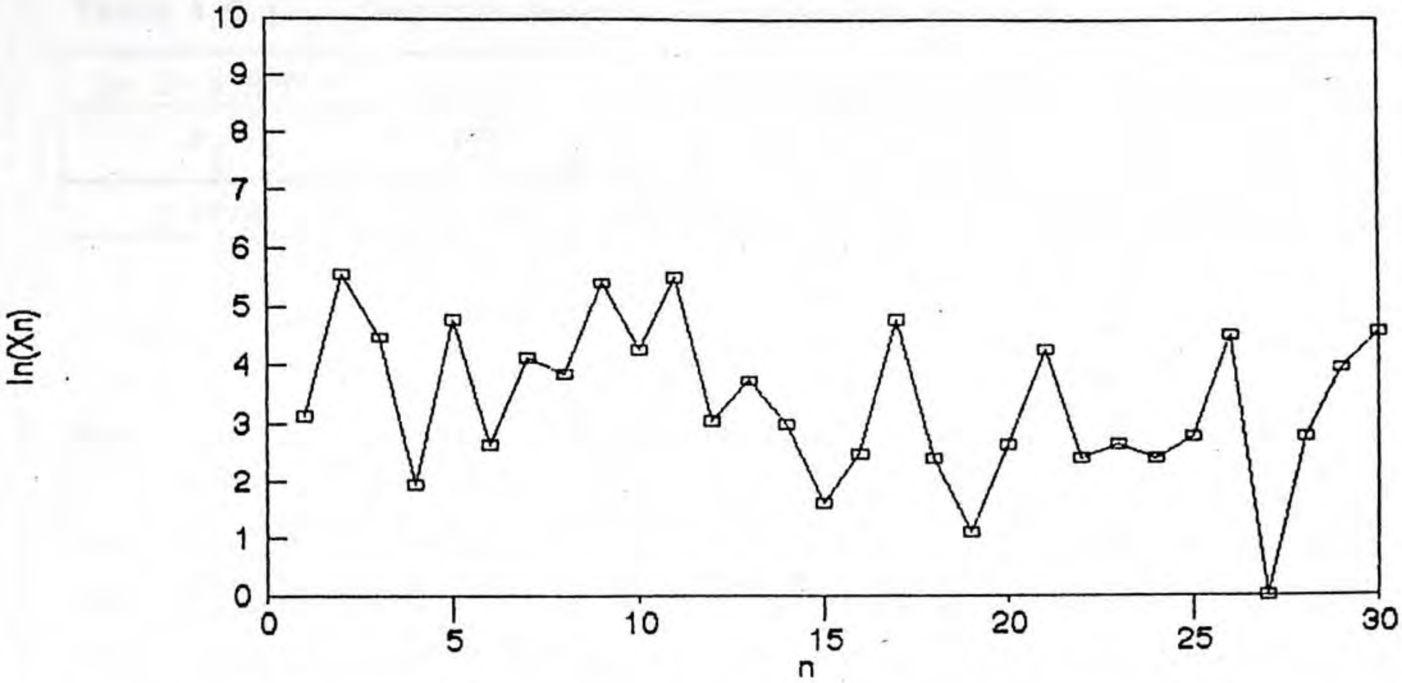
Is it a HPP ?	Is it a GP ?				Is a=1 ?
P_u	P_T^U	P_D^U	P_T^V	P_D^V	P_t
0.0274	0.6633	0.3865	0.8277	1.0000	0.0840

Table 4.4.2 : Test for Distribution for Air2 data.

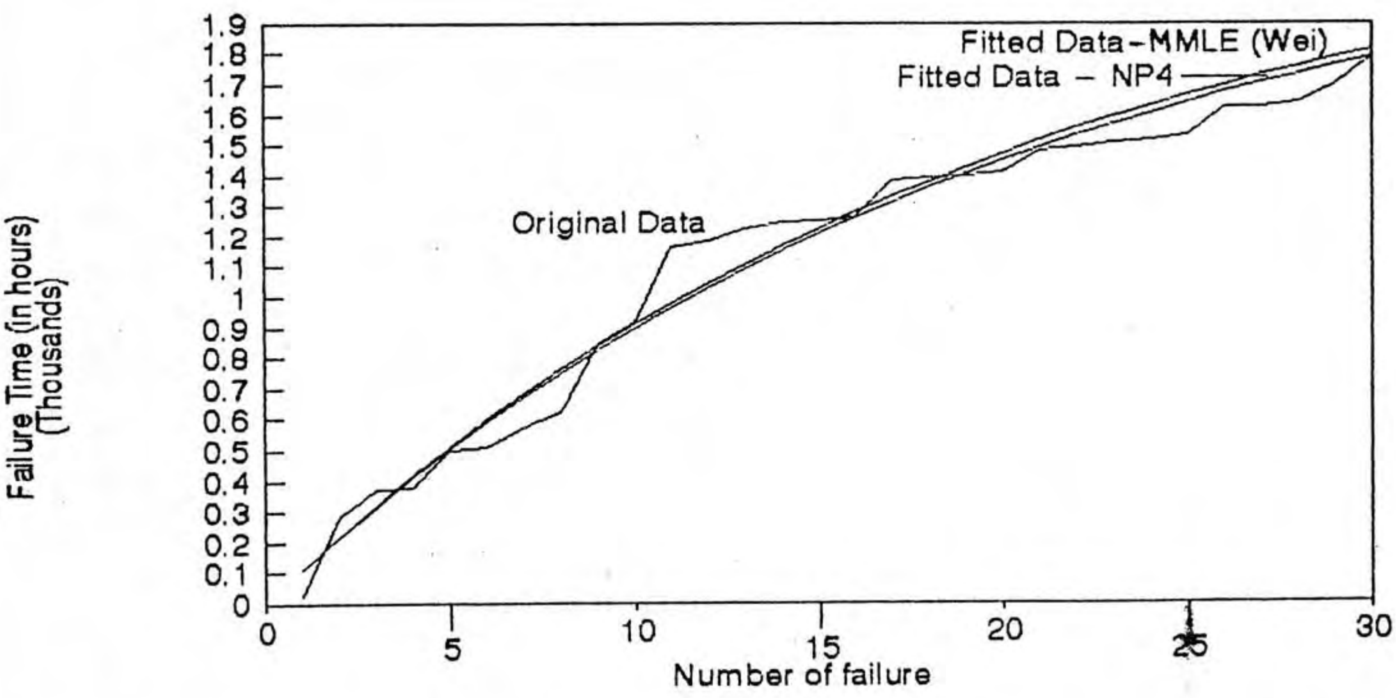
Method	KS-Test							F-Test	χ^2 GOF-Test (DF)
	W/W^*	U/U^*	A/A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D/D^*$	$\sqrt{N}V/V^*$		
MLE_E	*.1558	*.1276	*.8539			*.9804	1.7907	.0525	.1490 (3)
$MMLE_E$	*.1557	*.1252	*.8555			*.9037	*1.6385	.0809	.2276 (3)
$MLE1_E$	⊙.2147	⊙.1531	⊙1.1864			1.2036	2.2335	.0096	.3916 (3)
MLE_G	*.1066	*.0971	*.6015						.1146 (2)
$MMLE_G$	*.1044	*.0942	*.5936						.0421 (2)
MME_G	*.0908	*.0855	*.5381						.0421 (2)
$MLE1_G$	*.1172	*.1039	*.6534						.0080 (2)
$MMLE_W$	*.0898	*.0847	*.5354	*.5176	*.6899	*.6899	*1.2075		.0821 (2)
MME_W	*.0989	*.0642	*.5716	*.5215	*.7240	*.7240	*1.2455		.0421 (2)
$MLE1_W$	*.1010	*.0963	*.5724	*.5011	*.8400	*.8400	*1.3412		.0080 (2)
MLE_L	*.0373	*.0367	*.3026			*.4989	*.9347		.1599 (2)
MME_L	*.0373	*.0367	*.3026			*.4989	*.9347		.1599 (2)
$MLE1_L$	*.0703	*.0696	*.4070			*.5834	*1.1685		.0421 (2)

Table 4.4.3 : Estimation of Parameters for Air2 data.

Method	$\hat{a}$	$\hat{\lambda}$ Value	MSE	Rank	MSER	$\hat{\sigma}^2$	$\hat{\alpha}$ or $\hat{\mu}$	$\hat{\beta}$ or $\hat{\tau}^2$
NP1	1.050087	85.886	4594.94	14	4.81%	5698		
NP2		109.255	4384.76	2	0.02%	14966		
NP3		107.184	4388.44	10	0.10%			
NP4		110.874	4383.91	1	-			
NP5		112.784	4385.19	4	0.03%			
NP6	1.000000	59.600	4995.17	15	13.9%	5167		
MLE_E	1.042363	101.087	4401.41	11	0.40%	10218		0.0098924
$MMLE_E$	1.050087	112.784	4385.19	4	0.03%	12720		0.0088665
$MLE1_E$	1.000000	59.600	4995.17	15	13.9%	3552		0.0167785
MLE_G	1.042363	101.087	4401.41	11	0.45%	11458	0.89180	0.0088221
$MMLE_G$	1.050087	112.784	4385.19	4	0.03%	14311	0.88887	0.0078811
MME_G	1.050087	112.784	4385.19	4	0.03%	14967	0.84990	0.0075356
$MLE1_G$	1.000000	59.600	4995.17	15	13.9%	4375	0.81191	0.0136227
$MMLE_W$	1.050087	112.535	4384.89	3	0.02%	15374	0.90889	0.0093003
MME_W	1.050087	112.784	4385.19	4	0.03%	14967	0.92281	0.0092061
$MLE1_W$	1.000000	59.262	4995.29	18	13.9%	4856	0.85359	0.0183105
MLE_L	1.050087	127.405	4476.91	13	2.12%	61110	4.06675	1.5612464
MME_L	1.050087	112.784	4385.19	4	0.03%	14967	4.33659	0.7777699
$MLE1_L$	1.000000	68.593	5076.04	19	15.8%	22106	3.35809	1.7401908



Graph 4.4.1 : Plot of $\ln X_n$ against n for Air2 data.



Graph 4.4.2 : Plot of Air2 data and their fitted values using NP4 and MMLE (Wei).

Table 4.5.1 : Test for Geometric Process for No4 data.

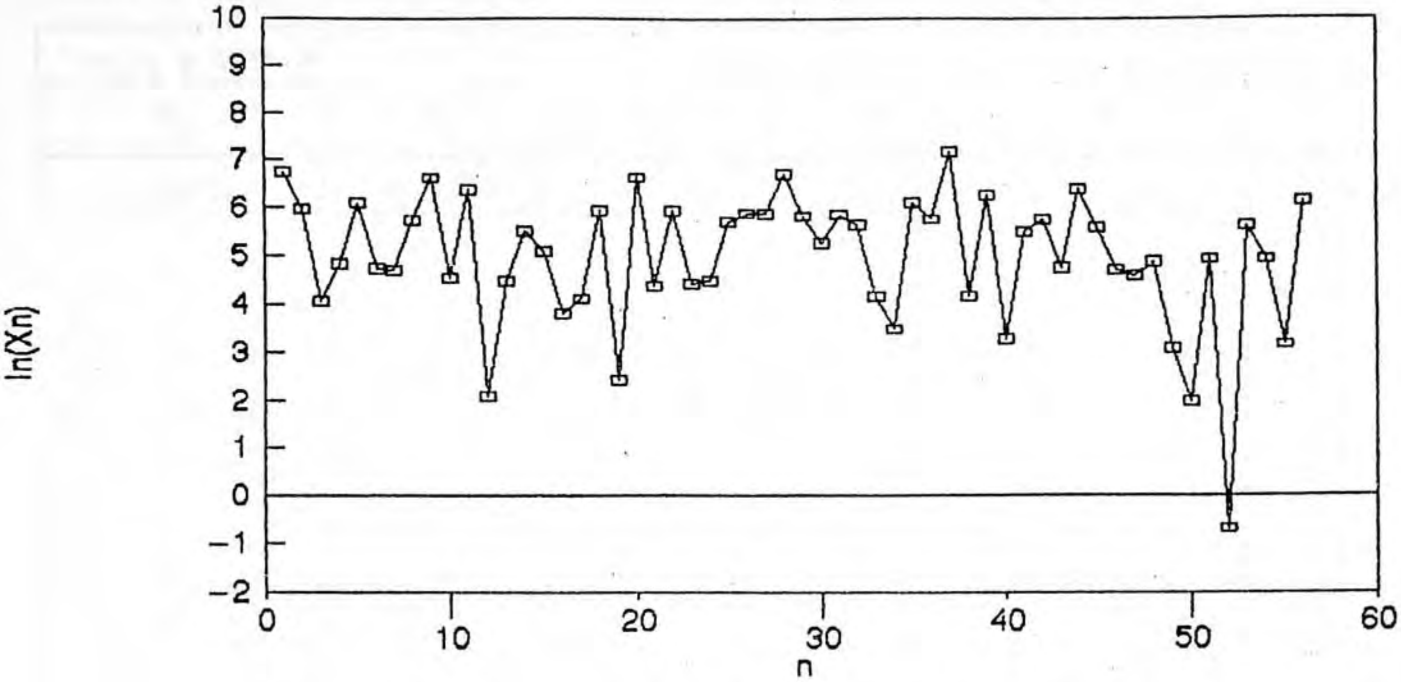
Is it a HPP ?	Is it a GP ?				Is a=1 ?
P_u	P_T^U	P_D^U	P_T^V	P_D^V	P_t
0.3176	0.7573	0.3346	0.0893	0.7477	0.1286

Table 4.5.2 : Test for Distribution for No4 data.

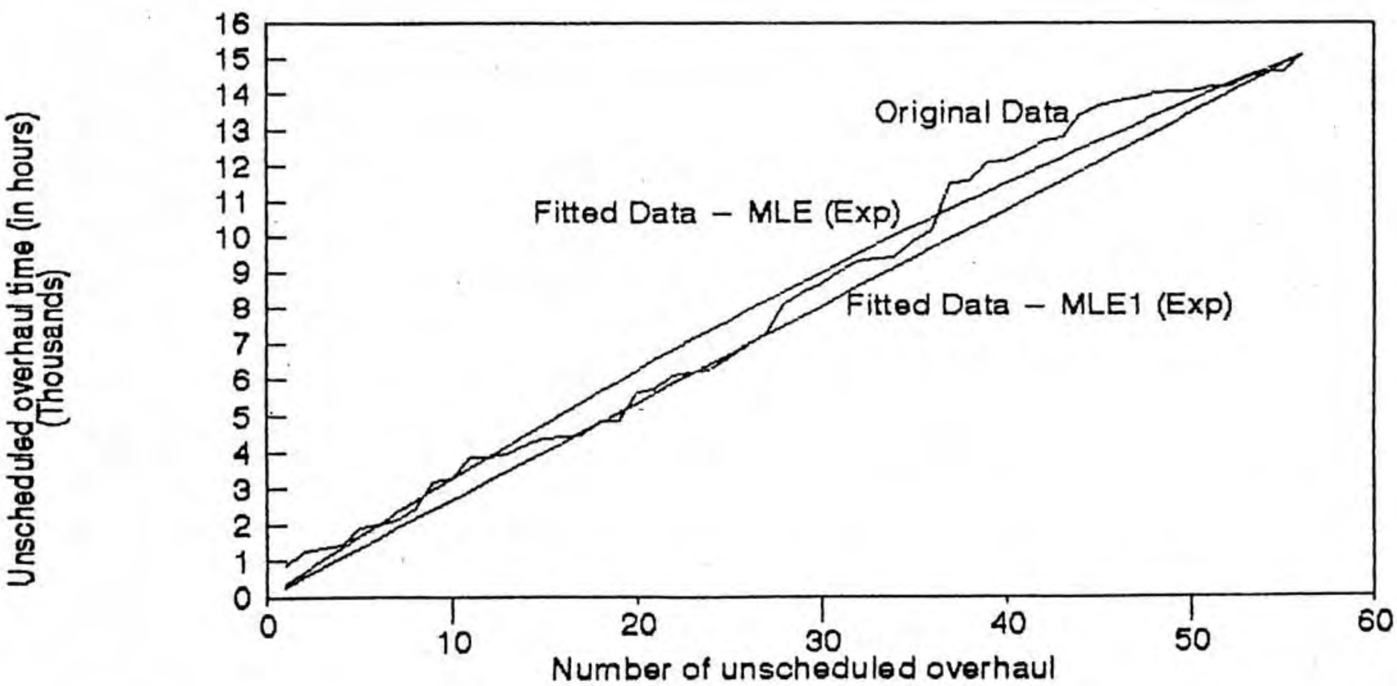
Method	KS-Test							F-Test	χ^2 GOF-Test (DF)
	W/W^*	U/U^*	A/A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D/D^*$	$\sqrt{N}V/V^*$		
MLE_E	*.0721	*.0721	*.4023			*.7140	*1.3097	.6345	.5494 (5)
$MMLE_E$	*.0723	*.0721	*.4157			*.8269	*1.5344	.8795	.0687 (6)
$MLE1_E$	*.0589	*.0589	*.3249			*.6044	*1.2105	.4045	.3313 (5)
MLE_G	*.0694	*.0669	*.3778						.5169 (4)
$MMLE_G$	*.0672	*.0659	*.3708						.2102 (5)
MME_G	*.0679	*.0677	*.3781						.3062 (5)
$MLE1_G$	*.0562	*.0538	*.2955						.2545 (5)
$MMLE_W$	*.0711	*.0702	*.3924	*.6380	*.7204	*.7504	*1.3585		.3062 (5)
MME_W	*.0715	*.0684	*.3988	*.6256	⊙.7615	*.7615	⊙1.3871		.3062 (5)
$MLE1_W$	.0582	.0570	.3101	.6201	.5325	.6201	1.1525		.2102 (5)
MLE_L	.2040	.1672	1.2738			⊙.9438	1.6268		.0266 (4)
MME_L	.2040	.1672	1.2738			⊙.9438	1.6268		.0266 (4)
$MLE1_L$	.2068	.1647	1.3120			⊙.9473	1.6155		.2397 (4)

Table 4.5.3 : Estimation of Parameters for No4 data.

Method	$\hat{a}$	$\hat{\lambda}$				$\hat{\sigma}^2$	$\hat{\alpha}$ or $\hat{\mu}$	$\hat{\beta}$ or $\hat{\tau}^2$
		Value	MSE	Rank	MSER			
NP1	1.018090	359.609	70202	16	4.04%	109493		
NP2		419.034	69424	7	2.89%	202376		
NP3		467.969	71112	17	5.39%			
NP4		422.633	69477	8	2.97%			
NP5		440.439	69904	9	3.60%			
NP6	1.000000	269.107	68629	3	1.71%	69877		
MLE_E	1.009062	341.572	67476	1	-	116671		0.0029276
$MMLE_E$	1.018090	440.439	69904	9	3.60%	193987		0.0022705
$MLE1_E$	1.000000	269.107	68629	3	1.71%	72419		0.0037160
MLE_G	1.009062	341.572	67476	1	-	123802	0.94241	0.0027590
$MMLE_G$	1.018090	440.439	69904	9	3.60%	208594	0.92997	0.0021115
MME_G	1.018090	440.439	69904	9	3.60%	202376	0.95855	0.0021763
$MLE1_G$	1.000000	269.107	68629	3	1.71%	77913	0.92948	0.0034539
$MMLE_W$	1.018090	440.637	69910	15	3.61%	208878	0.96432	0.0023062
MME_W	1.018090	440.439	69904	9	3.60%	202376	0.97912	0.0022914
$MLE1_W$	1.000000	269.268	68629	3	1.71%	76971	0.97069	0.0037626
MLE_L	1.018090	610.099	87932	19	30.3%	2126119	5.46167	1.9038955
MME_L	1.018090	440.439	69904	9	3.60%	202376	5.73050	0.7145407
$MLE1_L$	1.000000	388.619	82910	18	22.9%	951401	4.96864	1.9878697



Graph 4.5.1 : Plot of $\ln X_n$ against n for No4 data.



Graph 4.5.2 : Plot of No4 data and their fitted values using MLE (Exp) and MLE1 (Exp).

Table 4.6.1 : Test for Geometric Process for Patients data.

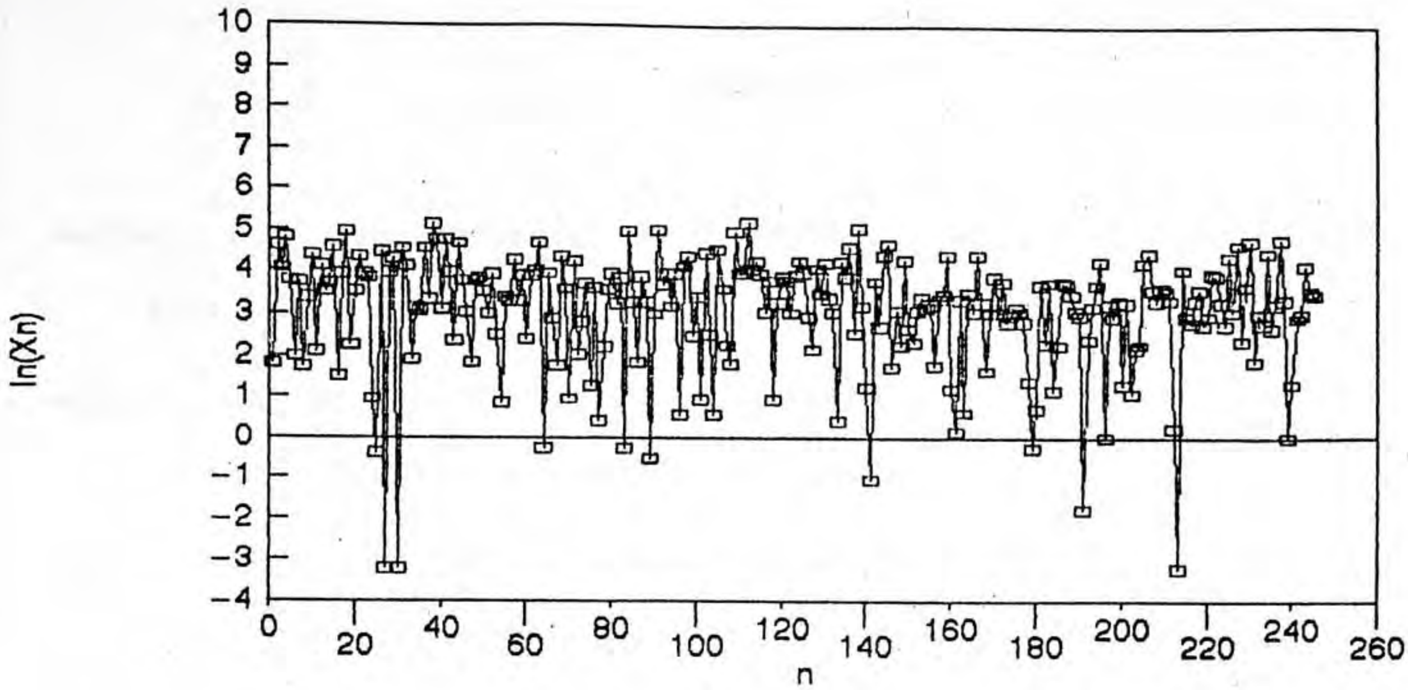
Is it a HPP ?	Is it a GP ?				Is a=1 ?
P_u	P_T^U	P_D^U	P_T^V	P_D^V	P_t
0.0096	0.3868	0.2743	0.2794	0.6394	0.2216

Table 4.6.2 : Test for Distribution for Patients data.

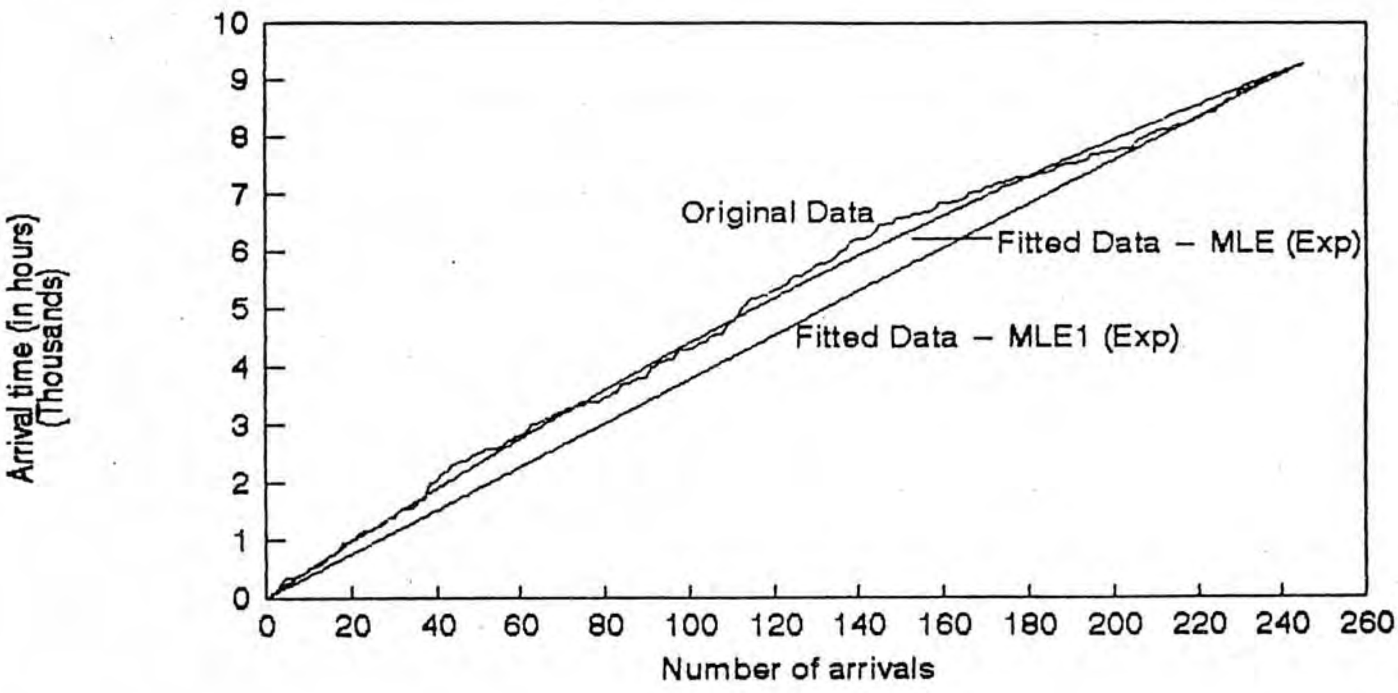
Method	KS-Test							F-Test	χ^2 GOF-Test (DF)
	W/W^*	U/U^*	A/A^*	$\sqrt{N}D^+$	$\sqrt{N}D^-$	$\sqrt{N}D/D^*$	$\sqrt{N}V/V^*$		
MLE_E	.3225	.2767	1.7214			1.3997	2.7422	.0515	.0084(18)
$MMLE_E$	.2852	.2432	⊙1.5583			1.3623	2.6676	.0828	.0037(18)
$MLE1_E$	*.1624	⊙.1417	*1.0385			⊙1.1197	2.1833	.3739	.0389(18)
MLE_G	.3796	.3061	1.8497						.0019(17)
$MMLE_G$	.3418	.2718	1.6943						.0041(17)
MME_G	.2011	.1984	2.4372						.0017(17)
$MLE1_G$	.2249	.1722	1.1567						.0251(17)
$MMLE_W$	.2680	.2344	1.5294	1.2980	*.6226	1.2980	1.9206		.0033(17)
MME_W	.1717	.1669	1.8198	.7959	1.0037	1.0037	1.7996		.0023(17)
$MLE1_W$	.1741	.1488	1.0664	1.1440	*.5469	1.1440	1.6909		.0151(17)
MLE_L	1.5975	1.3076	8.6950			2.6708	4.0518		.0000(11)
MME_L	1.5975	1.3076	8.6950			2.6708	4.0518		.0000(11)
$MLE1_L$	1.3917	1.1196	7.8005			2.4844	3.7781		.0000(11)

Table 4.6.3 : Estimation of Parameters for Patients data.

Method	$\hat{a}$	$\hat{\lambda}$ Value	MSE	Rank	MSER	$\hat{\sigma}^2$	$\hat{\alpha}$ or $\hat{\mu}$	$\hat{\beta}$ or $\hat{\tau}^2$
NP1	1.001611	38.4986	1217.03	17	3.72%	1302		
NP2		40.9936	1195.26	12	1.87%	1687		
NP3		51.0524	1194.99	11	1.84%			
NP4		45.7685	1177.65	3	0.37%			
NP5		45.5089	1177.80	4	0.38%			
NP6	1.000000	37.8505	1213.56	13	3.43%	1219		
MLE_E	1.002269	49.2490	1173.36	1	-	2425		0.0203050
$MMLE_E$	1.001611	45.5089	1177.80	4	0.38%	2071		0.0219737
$MLE1_E$	1.000000	37.8505	1213.56	13	3.43%	1433		0.0264197
MLE_G	1.002269	49.2490	1173.36	1	-	2531	0.95822	0.0194566
$MMLE_G$	1.001611	45.5089	1177.80	4	0.38%	2165	0.95664	0.0210209
MME_G	1.001611	45.5089	1177.80	4	0.38%	1687	1.22733	0.0269691
$MLE1_G$	1.000000	37.8505	1213.56	13	3.43%	1524	0.93984	0.0248304
$MMLE_W$	1.001611	45.4913	1177.81	10	3.79%	2028	1.01008	0.0218904
MME_W	1.001611	45.5089	1177.80	4	0.38%	1687	1.10947	0.0211436
$MLE1_W$	1.000000	37.8588	1213.56	13	3.43%	1451	0.99389	0.0264831
MLE_L	1.001611	70.8275	1603.17	18	36.6%	35860	3.21134	2.0978105
MME_L	1.001611	45.5089	1177.80	4	0.38%	1687	3.51993	0.5959610
$MLE1_L$	1.000000	58.5748	1643.05	19	40.0%	24891	3.01492	2.1107770



Graph 4.6.1 : Plot of $\ln X_n$ against n for Patients data.



Graph 4.6.2 : Plot of Patients data and their fitted values using MLE (Exp) and MLE1 (Exp).

REFERENCE

Andrews, D. F. and Herzberg, A. M. (1985), *Data,* Springer-Verlag, New York.

Anderson, T. W. and Darling, D. A. (1954), *A Test of Goodness-of-fit,* J. Amer. Statist. Assoc., 49, 765-769.

Ascher, H. and Feingold, H. (1969), *'Bad-As-Old' Analysis of System Failure Data,* Annuals of Assurance Sciences, Gordan and Breach, New York, 49-62.

Ascher, H. and Feingold, H. (1984), *Repairable Systems Reliability,* Marcel dekker, New York.

Bain, L. J. and Engelhardt, M. (1991), *Statistical Analysis of Reliability and Life-testing Models, Theory and Methods,* Marcel Dekker, New York.

Cao, Jinhua and Cheung, Kan (1986), *Introduction to Mathematical Theory of Reliability,* Science Press. (in Chinese)

Chandra, M., Singpurwalla, N. D., and Stephens, M. A. (1981), *Kolmogorov Statistics for Tests of Fit for the Extreme-value and Weibull Distributions,* J. Amer. Statist. Assoc. 76, 729-731.

Cohen, A. C. (1973), *The Reflected Weibull Distribution,* Technometrics, 15, 867-873.

Cohen, A. C. and Whitten, B. J. (1988), *Parameter Estimation in Reliability and Life Span Models,* Marcel dekker, New York.

Cox, D. R. and Lewis, P. A. (1966), *The Statistical Analysis of Series of Events,* Mathuen, London.

D'Agostino, R. B. and Stephens, M. A. (eds) (1986), *Goodness-of-fit Techniques,* Marcel dekker, New York.

Durbin, J. (1975), *Kolmogorov-Smirnov Tests when Parameters are Estimated with Applications to Tests of Exponentiality and Tests on Spacings,* Biometrika 62, 5-22.

Epstein, B. (1958), *Exponential Distribution and its Role in Life Testing,* Industrial Quality Control, 15, 4-6.

Jarrett, R. G. (1979), *A Note on the Intervals between Coal-mining Disasters,* Biometrika, 66, 191-193.

Johnson, N. L. and Kotz, S. (1970), *Continuous Univariate Distributions,* I. Houghton Mifflin, Boston.

Kao, J. H. K. (1959), *A Graphical Estimation of Mixed Weibull Parameters in Life Testing Election Tubes,* Technometrics 1, 389-407.

Kolmogorov, A. N. (1933), *Sulla determinazione empirica di una legge di distibuziane,* Giorna. Ist. Attuari., 4, 83-91.

Kuiper, N. H. (1960), *Tests concerning Random Points on a Circle,* Proc. Koninkl. Neder. Akad. van. Wetenschappen, A 63, 38-47.

Lam, Yeh (1988a), *A Note on the Optimal Replacement Problem,* Adv. Appl. Prob., 20, 479-482.

Lam, Yeh (1988b), *Geometric Process and Replacement Problem,* Acta Math. Appl. Sinica 4, 366-377.

Lam, Yeh (1992a), *Optimal Geometric Process Replacement Model,* Acta Math. Appl. Sinica 8, 73-81.

Lam, Yeh (1992b), *Nonparametric Inference for Geometric Processes,* Commun. Statist.-Theor. Meth., 21 (in press).

Lee, L. (1980a), *Testing Adequacy of the Weibull and Log Linear Rate Models for a Poisson Process,* Technometrics, 22, 195-199.

Lee, L. (1980b), *Comparing Rates of Several Independent Weibull Processes,* Technometrics, 22, 427-430.

Lieblein, J. and Zelen, M. (1956), *Statistical Investigation of the Fatigue Life of Deep Groove Ball Bearings,* J. of Res. Natl. Bur, Std., 57, 273-315.

Lilliefors, H. W. (1969), *On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown,* J. Amer. Statist. Assoc., 64, 387-389.

Lockhardt R. A. and Stephens, M. A. (1985b), *Goodness-of-fit Tests for the Gamma Distribution,* Technical Report, Department of Mathematics and Statistics, Simon Fraser University.

Maguire, B. A., Pearson, E. S. and Wynn, A. H. A. (1952), *The Time Intervals between Industrial Accidents,* Biometrika, 39, 168-180.

Margolin, B. H. and Maurer, W. (1976), *Tests of the Kolmogorov-Smirnov type for Exponential Data with Unknown Scale, and Related Problems,* Biometrika 63, 149-160.

Pearson, E. S. and Hartley, H. O. (1972), *Biometrika Tables for Statisticans,* Vol. 2, New York: Cambridge University Press.

Proschan, F. (1963), *Theoretical Explanation of Observed Decreasing Failure Rate,* Technometrics, 5, 375-383.

Sinha, S. K. (1986), *Reliability and Life Testing,* John Wiley & Sons, New York.

Stephens, M. A. (1974b), *EDF Statistics for Goodness-of-fit and some Comparisons,* J. Amer. Statist. Assoc. 69, 730-737.

Stephens, M. A. (1976a), *Asymptotic Results for Goodness-of-fit Statistics with Unknown Parameters,* Ann. Statist. 4, 357-369.

Stephens, M. A. (1977), *Goodness-of-fit for the Extreme Value Distribution,* Biometrika 64, 583-588.

Watson, G. S. (1961), *Goodness-of-fit Tests on a Circle,* Biometrika 48, 109-114.

Weibull, W. (1939), *A Statistical Theory of the Strength of Materials,* Ing. Vetenskps Akad. Handl. 151, 1-45.

Weibull, W. (1951), *A Statistical Distribution Function of Wide Applicability,* J. Appl. Mech. 18, 293-297.